1927/1928

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

S. LEFSCHETZ J. H. M. WEDDERBURN

WITH THE COÖPERATION OF

ORMOND STONE W. B. FORD
D. C. GILLESPIE W. L. HART
E. R. HEDRICK EINAR HILLE
R. E. LANGER A. PELL-WHEELER

PUBLISHED BY THE

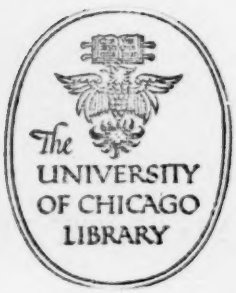
PRINCETON UNIVERSITY PRESS

SECOND SERIES, VOL. 29

PRINCETON, N. J.

1928

QA1
.A65
Ser. 2
v. 29
Cop. 2



Gift - G. A. Bliss

LÜTCKE & WULFF, HAMBURG, GERMANY

715

INDEX

PAGE

ALTSHILLER-COURT, N., On four mutually orthogonal circles 369

BALLANTINE, J. P., A generalization of the calculus of finite differences to include the differential calculus 14

BELL, E. T., Partition polynomials 38

BELL, J. L., Chains of congruences for the numerators and denominators of the Bernoulli numbers 106

BENDER, H. A., A determination of groups of order p^5 61

BENDER, H. A., On groups of order p^m, p being an odd prime number, which contain an abelian subgroup of order p^{m-1} 88

BOGOLIOUBOFF, N., and N. KRYLOFF, On Rayleigh's principle in the theory of differential equations and on Euler's method in calculus of variations 255

BROWNE, E. T., Involutions that belong to a linear class 483

CAMPBELL, A. D., The discriminant of the m-ary quadratic in the Galois fields of order 2^n 395

CHOU, P. Y., A new derivation of the Lorentz transformation 433

DAVIS, H. T., On the factoring of Fredholm minors 334

DORROH, J. L., Concerning a set of metrical hypothesis for geometry . 229

DOUGLAS, J., The general geometry of paths 143

DOUGLAS, J., A method of numerical solution of the problem of Plateau . 180

FLEXNER, W. W., A general theorem on the expression of a determinant in terms of its sub-determinants 373

FORD, L. R., On motions which satisfy Kepler's first and second laws . 549

FRANKLIN, P., The canonical form of a one-parameter group 113

GARVER, R., The Tschirnhaus transformation 319

GLENN, O. E., On the generalization of the algebra of the theory of numbers. The invariancy of infinite series 459

GUGGENBÜHL, L., An integral equation with an associated integral condition 21

HASKELL, R. N., A Note on Stieltjes integrals 543

HILLE, E., A class of functional equations 215

HILTON, H., On two types of plane rational curve 194

HOPKINS, C., Non-abelian groups whose groups of isomorphisms are abelian 508

JAMES, G., An integration method of summing series 79

JANCZEWSKY, S. A., Oscillation theorems for the differential boundary value problems of the fourth order 521

KERÉKJÁRTÓ, B. DE, On a geometrical theory of continuous groups. II. Euclidean and hyperbolic groups of three-dimensional space 169

PAGE
KNEBELMAN, M. S., Collineations of projectively related affine connections 389
KRYLOFF, N., See BOGOLIOUBOFF.
LEFSCHETZ, S., Closed point sets on a manifold 232
MACDUFFEE, C. C., A correspondence between matrices and quadratic ideals 199
MERRIMAN, G. M., A set of necessary and sufficient conditions for the Cesàro summability of double series 343
MILLER, G. A., Definitions of abstract groups 223
MURRAY, F. H., A class of real quadratic forms in infinitely many variables 123
OGG, F. C., Certain configurations on cubics 355
RANUM, A., Twisted curves classified as to their osculating and quasi-osculating spheres 445
RIDER, P. R., Minimizing two types of definite integral 189
SEELY, C., Kernel's of positive type 313
STONE, M. H., Developments in Hermite polynomials. 1
SUN, D., Osculating derivative of a ruled surface 95
TRJITZINSKY, W. J., Relations satisfied by coefficients of periodic solutions 17
TRJITZINSKY, W. J., Representation of functions determined by their initial values 73
WALSH, C. M., Fermat's Note XLV 412
WEIDA, F. M., On various conceptions of correlation 276
WEISNER, L., Quadratic fields in which cyclotomic polynomials are reducible 377
WEISNER, L., Euclidean invariants of plane algebraic curves 493
WEISNER, L., On m-dimensional cross-ratios 440
WHYBURN, G. T., Concerning the complementary domains of continua . 399
WILSON, W. A., On irreducible cuts of the plane between two points . 382
YOUNG, J. W., A new formulation for general algebra 47

ERRATA. Volume 28.

P. 6 (10a): For "12 —" read "12 +".
„ 7 (14): „ "$52+(c-\mu)$" read "$52-2a+(c-\mu)$".
„ 12 (20): „ "$n>9$" read "$n \geqq 9$".
„ 47, last line, should read: $(-\alpha'_{\nu-2}\mu^2+\beta'_{\nu-2}\mu+\gamma''_{\nu-2})/R'_{\nu-2}=\varrho'_{\nu-2}$.
Hence

$$\psi_{\nu-1}(x)=R'_{\nu-1}x^3-(3\gamma'_{\nu-1}-2A_2\alpha'_{\nu-1})x^2+(3\gamma''_{\nu-2}+2A_2\alpha'_{\nu-2})x-R'_{\nu-2}$$

and therefore

$$-\frac{R_n}{R'_{\nu-1}}=\frac{\psi(1)}{R'_{\nu-1}}\to 1-4+5-2=0.$$

From this point on the discussion must proceed on other lines in order to prove that $R_n\to\infty$ as well as to give conditions for ω_n being limited.

P. 480 (4g): For "p. 232" read "p. 233".
„ 487 - 488: „ "$3\pi_\nu$" read "π_ν" in the tabular headings.
„ 490 (5.2): „ "p_2" read "$3p_2$".

DEVELOPMENTS IN HERMITE POLYNOMIALS.*

By M. H. Stone.

In the present paper we wish to give a brief account of an elementary treatment of developments in Hermite polynomials similar to one of developments in Legendre polynomials which we have given previously.† Because of the important application of such developments in the Gram-Charlier theory of frequency curves a discussion of this character seems desirable. A bibliography of the literature dealing with the Hermite polynomials is available in a recent article by Hille, to which we refer the reader for information concerning more elaborate and thorough discussions.‡ We shall follow the notation of Hille's paper in the present one.

I. Definition and Properties of the Hermite Polynomials.

The Hermite polynomial $H_n(x)$, $n = 0, 1, 2, \cdots$, is the polynomial in x of degree n defined by the development

$$G(x, r) \equiv e^{2xr-r^2} = \sum_{n=0}^{\infty} H_n(x)\, r^n/n!, \quad -\infty < x < +\infty, \quad |r| < +\infty.$$

If we expand the identities

$$\frac{\partial G}{\partial x} \equiv 2rG,$$

$$\frac{\partial^2 G}{\partial x^2} - 2x\frac{\partial G}{\partial x} + 2r\frac{\partial G}{\partial r} \equiv 0,$$

tn power series about $r = 0$ and equate the coefficients of like powers of r, we obtain respectively

$$H_n'(x) = 2n\, H_{n-1}(x), \qquad n = 1, 2, 3, \cdots, \tag{1}$$

$$H_n''(x) - 2x\, H_n'(x) + 2n\, H_n(x) = 0, \quad n = 0, 1, 2, \cdots. \tag{2}$$

If we eliminate the derivatives in (2) by the use of (1) we find the recurrence relation

$$H_n(x) - 2x\, H_{n-1}(x) + 2(n-1)\, H_{n-2}(x) = 0, \quad n = 2, 3, 4, \cdots, \tag{3}$$

* Received March 30, 1927.

† Stone, Annals of Math., vol. 27 (1926), pp. 315–329.

‡ Hille, Annals of Math., vol. 27 (1926), pp. 427–464.

from which we can construct the Hermite polynomials starting from $H_0 = 1$, $H_1 = 2x$. From the definition of $H_n(x)$ as the coefficient of $r^n/n!$ in the power series for $G(x, r)$ we see that

$$(4)\quad H_n(x) = \frac{\partial^n}{\partial r^n} G(x, r)\Big]_{r=0} = e^{x^2} \frac{\partial^n}{\partial r^n} e^{-(x-r)^2}\Big]_{r=0} = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}.$$

We may also, by comparison with the familiar series for $G(0, r)$, show that

$$(5)\quad H_{2m}(0) = (-1)^m \frac{(2m)!}{m!}, \quad H_{2m+1}(0) = 0,$$

and hence, by reference to (1), that

$$(6)\quad H'_{2m}(0) = 0, \quad H'_{2m+1}(0) = (-1)^m \frac{2(2m+1)!}{m!}.$$

Finally we deduce from the identity $G(-x, r) \equiv G(x, -r)$ the relation

$$(7)\quad H_n(-x) = (-1)^n H_n(x).$$

We shall now prove

THEOREM I. *The Hermite polynomials $H_n(x)$ satisfy the equations*

$$\int_{-\infty}^{+\infty} e^{-x^2} H_n(x) H_m(x)\, dx = 0, \qquad n \neq m,$$
$$\int_{-\infty}^{+\infty} e^{-x^2} H_n^2(x)\, dx = \sqrt{\pi}\, 2^n n!,$$

for $m, n = 0, 1, 2, \cdots$.

Since m and n may be interchanged in our reasoning, we may assume m less than or equal to n. Thus from (4)

$$\int_{-\infty}^{+\infty} e^{-x^2} H_n H_m\, dx = (-1)^n \int_{-\infty}^{+\infty} H_m \frac{d^n}{dx^n} e^{-x^2}\, dx = \int_{-\infty}^{+\infty} e^{-x^2} H_m^{(n)}\, dx,$$

by repeated integrations by parts. If m is less than n this last integral vanishes. If m and n are equal it has the value

$$n!\, A_n \int_{-\infty}^{+\infty} e^{-x^2}\, dx = n!\, A_n \sqrt{\pi}$$

where A_n is the coefficient of the highest power of x in H_n. From (1) we see that $A_n = 2A_{n-1}$, whence $A_n = 2^n$, $A_0 = 2^n$. The theorem follows at once.

When we come to the investigation of expansions in terms of the polynomials $H_n(x)$, we shall require certain inequalities for them. In order to establish these we need

LEMMA I. *There exists a positive constant H such that*

$$|H_n(0)| < H 2^{n/2} (n!)^{1/2} n^{-1/4}, \qquad |H_n'(0)| < H 2^{n/2} (n!)^{1/2} n^{1/4},$$

for $n = 1, 2, 3, \cdots$.

The DeMoivre-Stirling formula $m! = C m^{m+1/2} e^{-m} (1 + \varepsilon_m)$ where $\varepsilon_m = o(1)$ is very easily proved when no attempt is made to evaluate the constant C. Using this formula together with (5) and (6), we find for $n = 2m$

$$\begin{aligned} |H_n(0) 2^{-n/2} (n!)^{-1/2} n^{1/4}| &= 2^{-m+1/4} [(2m)!]^{1/2} (m!)^{-1} m^{-1/4} \\ &= C^{-1/2} 2^{1/2} (1+\varepsilon_{2m})^{1/2} (1+\varepsilon_m)^{-1} \\ &< H, \end{aligned}$$

and for $n = 2m+1$

$$\begin{aligned} |H_n'(0) 2^{-n/2} (n!)^{-1/2} n^{-1/4}| &= 2^{-m+1/4} [(2m+1)!]^{1/2} (m!)^{-1} \left(m + \frac{1}{2}\right)^{-1/4} \\ &= C^{-1/2} 2 \left(1 + \frac{1}{2m}\right)^{m+1/2} e^{-1/2} (1+\varepsilon_{2m+1})^{1/2} (1+\varepsilon_m)^{-1} \\ &< H. \end{aligned}$$

This completes the lemma.

We now proceed to

THEOREM II. *The Hermite polynomial $H_n(x)$ is such that*

$$|H_n(x)| < K 2^{n/2} (n!)^{1/2} n^{-1/4} (1 + |x|^{5/2}) e^{x^2/2}, \quad -\infty < x < +\infty,$$

for $n = 1, 2, 3, \cdots$, *where K is a positive constant independent of n and x.*

If we express $H_n(x)$ in the form $e^{x^2/2} \varphi_n(x)$ and substitute in (3), we find that φ_n satisfies the differential equation

$$\varphi_n'' + 2n \varphi_n = (x^2 - 1) \varphi_n \tag{8}$$

with the auxiliary conditions

$$\varphi_n(0) = H_n(0), \qquad \varphi_n'(0) = H_n'(0).$$

From φ_n we define two functions, θ_n and u_n. The first of these is given as

$$\theta_n = \frac{1}{\sqrt{2n}} \int_0^x \sin \sqrt{2n} (x-y) (1-y^2) \varphi_n(y) \, dy;$$

it may be verified directly that θ_n satisfies the differential equation

$$\theta_n'' + 2n \theta_n = (1 - x^2) \varphi_n \tag{9}$$

together with the boundary conditions

$$\theta_n(0) = 0, \qquad \theta_n'(0) = 0.$$

The function u_n is defined as the sum of φ_n and θ_n. By adding (8) and (9) we see that this function must satisfy the differential equation

$$u_n'' + 2n\,u_n = 0; \tag{10}$$

furthermore $u_n(0) = H_n(0)$, $u_n'(0) = H_n'(0)$. The only solution of (10) satisfying these conditions is found by quadratures to be

$$u_n = H_n(0)\cos\sqrt{2n}\,x + \frac{H_n'(0)}{\sqrt{2n}}\sin\sqrt{2n}\,x.$$

We now estimate the magnitudes of $|u_n|$ and $|\theta_n|$. From Lemma I,

$$|u_n| < 2H\,2^{n/2}(n!)^{1/2}\,n^{-1/4}, \qquad -\infty < x < \infty.$$

By the use of Schwarz's inequality we find that

$$\begin{aligned}|\theta_n| &\leqq \frac{1}{\sqrt{2n}}\left|\int_0^x |1-y^2|\,|\varphi_n(y)|\,dy\right| \\ &\leqq \frac{1}{\sqrt{2n}}\left(\int_0^x (1-y^2)^2\,dy\cdot\int_0^x \varphi_n^2(y)\,dy\right)^{1/2} \\ &\leqq \frac{1}{\sqrt{2n}}\,|p(x)|^{1/2}\left(\int_{-\infty}^{+\infty}\varphi_n^2(y)\,dy\right)^{1/2}\end{aligned}$$

where $p(x)$ is a polynomial of the fifth degree in x whose coefficients are independent of n. The integral may be evaluated by expressing φ_n in terms of H_n and referring to Theorem I. Thus

$$|\theta_n| < k(1+|x|^{5/2})\,2^{n/2}(n!)\,n^{-1/2}, \qquad -\infty < x < \infty$$

where k is a positive constant independent of n. Finally

$$|H_n(x)| = e^{x^2/2}\,|\varphi_n(x)| \leqq e^{x^2/2}(|u_n|+|\theta_n|)$$

so that

$$|H_n(x)| < K\,2^{n/2}(n!)^{1/2}\,n^{-1/4}(1+|x|^{5/2})\,e^{x^2/2}$$

where K is a positive number independent of n and x.

II. Expansions in Hermite Polynomials.

We shall now investigate the possibility of representing an "arbitrary" function by the series

$$f(x) \sim c_0 H_0(x) + c_1 H_1(x) + \cdots + c_n H_n(x) + \cdots,$$

$$c_n = \frac{1}{2^n n!\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-x^2} f(x)\,H_n(x)\,dx,$$

and the series

$$F(x) \sim C_0 e^{-x^2} H_0(x) + C_1 e^{-x^2} H_1(x) + \cdots + C_n e^{-x^2} H_n(x) + \cdots,$$

$$C_n = \frac{1}{2^n n! \sqrt{\pi}} \int_{-\infty}^{+\infty} F(x) H_n(x)\, dx.$$

The second series, which may be called the Gram-Charlier series, is related to the first by the fact that one can be obtained from the other by the substitution $F \equiv e^{-x^2} f$. For the coefficients in the two series we prove the following uniqueness theorem.

THEOREM III. *If $\varphi(x)$ is a continuous function which for large values of $|x|$ is of the form $O(e^{-kx^2})$ where k is a positive constant, and if*

$$\int_{-\infty}^{+\infty} \varphi(x) H_n(x)\, dx = 0, \qquad n = 0, 1, 2, \cdots,$$

then $\varphi(x)$ is identically zero.

If $\varphi(x) \not\equiv 0$ for $x = \xi$ we may assume that $\varphi(\xi) > 0$. We set $\psi(x) = \varphi(x + \xi)$ so that

$$\psi(0) = \varphi(\xi) > 0; \quad \psi(x) = O(e^{-lx^2}), \quad 0 < l < k;$$

and

$$\int_{-\infty}^{+\infty} \psi(x) x^n\, dx = \int_{-\infty}^{+\infty} \varphi(x) (x - \xi)^n\, dx = 0,$$

the polynomial $(x - \xi)^n$ being representable as a linear combination of the polynomials $H_0, H_1, \cdots, H_{n-1}, H_n$. Now by Taylor's theorem with remainder

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^{n-1}}{n-1!} + R_n(z),$$

$$R_n(z) = \frac{e^{\theta z}}{n!} z^n, \quad 0 < \theta < 1, \quad -\infty < z < +\infty.$$

On replacing z by $-\alpha^2 x^2$ we obtain

$$e^{-\alpha^2 x^2} = \sum_0^{n-1} \frac{(-1)^k \alpha^{2k} x^{2k}}{k!} + r_n(x),$$

$$r_n(x) = (-1)^n \frac{e^{-\theta \alpha^2 x^2}}{n!} \alpha^{2n} x^{2n}.$$

Consequently,

$$I(\alpha) = \int_{-\infty}^{+\infty} \psi(x) e^{-\alpha^2 x^2}\, dx = \int_{-\infty}^{+\infty} \psi r_n\, dx = O\left(\frac{\alpha^{2n}}{n!} \int_{-\infty}^{+\infty} e^{-lx^2} x^{2n}\, dx\right)$$

$$= O\left(\frac{\alpha^{2n} (2n)!}{2^n l^n (n!)^2}\right) \qquad = O\left(n^{-1/2} \frac{2^n \alpha^{2n}}{l^n}\right)$$

$$= o(1), \qquad 0 \leqq \alpha^2 \leqq l/2.$$

The function $I(\alpha)$ is analytic in α when the real part of $-\alpha^2-l^2$ is negative; it vanishes on a segment of the real axis and must therefore vanish identically. The equation $\alpha I(\alpha)=0$, $-\infty<\alpha<+\infty$, is incompatible with the inequality $\psi(0)>0$, as we shall now show. Because of the continuity of $\psi(x)$ we can determine ε so that $\psi(x)\geqq m>0$, $-\varepsilon\leqq x\leqq\varepsilon$. Thus

$$\int_{-\infty}^{-\varepsilon}\psi(x)\,\alpha e^{-\alpha^2x^2}\,dx = O(\alpha e^{-\alpha^2\varepsilon^2}) = o(1),$$

$$\int_{+\varepsilon}^{+\infty}\psi(x)\,\alpha e^{-\alpha^2x^2}\,dx = O(\alpha e^{-\alpha^2\varepsilon^2}) = o(1),$$

as $\alpha\to+\infty$; and

$$\int_{-\varepsilon}^{+\varepsilon}\psi(x)\,\alpha e^{-\alpha^2x^2}\,dx \geqq m\int_{-\varepsilon}^{+\varepsilon}\alpha e^{-\alpha^2x^2}\,dx = m\int_{-\varepsilon\alpha}^{+\varepsilon\alpha}e^{-x^2}\,dx = m\sqrt{\pi}+o(1).$$

We find therefore that $0=\alpha I(\alpha)\geqq m\sqrt{\pi}+o(1)$, whence we must have $m\leqq 0$. This contradiction shows that $\psi(0)=\varphi(\xi)=0$, $-\infty<\xi<+\infty$.

Concerning the magnitude of the coefficients c_n we can obtain certain information in case that the integral $\int_{-\infty}^{+\infty}e^{-x^2}f^2\,dx$ exists in the sense of either Riemann or Lebesgue.

THEOREM IV. *If the integral* $\int_{-\infty}^{+\infty}e^{-x^2}f^2\,dx$ *exists, the quantities*

$$c_n = \frac{1}{2^n n!\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-x^2}fH_n\,dx$$

are such that the infinite series

$$\sum_{n=0}^{\infty}2^n n!\,c_n^2$$

converges.

We see that

$$\int_{-\infty}^{+\infty}e^{-x^2}\Big(f-\sum_0^N c_n H_n\Big)^2 dx \geqq 0.$$

On expanding the integral on the left we obtain

$$\int_{-\infty}^{+\infty}e^{-x^2}f^2\,dx - \sum_0^N\sqrt{\pi}\,2^n n!\,c_n^2\geqq 0,$$

$$\int_{-\infty}^{+\infty}e^{-x^2}f^2\,dx \geqq \sum_0^N\sqrt{\pi}\,2^n n!\,c_n^2.$$

If we allow N to become infinite the theorem is established.

It is now possible to prove the first convergence theorem.

THEOREM V. *If $f(x)$ can be written in the form*

$$f(x) = f(0) + \int_0^x f'\,dx, \qquad -\infty < x < +\infty,$$

where

$$\int_{-\infty}^{+\infty} e^{-x^2} g^2\,dx, \qquad g = -f' + 2xf,$$

exists and where

$$f(x) = O(e^{kx^2}), \qquad k < 1,$$

then the series

$$\sum_0^\infty c_n\, e^{-x^2} H_n(x), \qquad c_n = \frac{1}{2^n\, n!\, \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} f H_n\,dx,$$

converges uniformly to $e^{-x^2} f(x)$, $-\infty < x < +\infty$; and the series

$$c_0 H_0(x) + c_1 H_1(x) + \cdots + c_n H_n(x) + \cdots$$

converges uniformly to $f(x)$ on any finite interval.

From formula (1) we see that

$$\begin{aligned} c_n &= \frac{1}{2^n\, n!\, \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} f H_n\,dx = \frac{1}{2^{n+1}(n+1)!\, \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} f H_{n+1}\,dx \\ &= \frac{1}{2^{n+1}(n+1)!\, \sqrt{\pi}} \left\{ [e^{-x^2} f H_{n+1}]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-x^2} g H_{n+1}\,dx \right\} \\ &= \frac{1}{2^{n+1}(n+1)!\, \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} g H_{n+1}\,dx = C_{n+1}, \end{aligned}$$

since the first term in the parenthesis vanishes at both limits under our restrictions on $f(x)$. By the use of Theorem II and the inequality $|ab| \leqq \frac{a^2+b^2}{2}$, we find for $n = 1, 2, 3, \cdots$

$$\begin{aligned} |c_n\, e^{-x^2} H_n(x)| &\leqq K e^{-x^2/2}(1 + |x|^{5/2})\, 2^{n/2} (n!)^{1/2}\, n^{-1/4}\, |C_{n+1}| \\ &\leqq K' e^{-lx^2/2} \{2^{(n+1)/2} \sqrt{(n+1)!}\, (n+1)^{-3/4}\, |C_{n+1}|\} \\ &\leqq K' e^{-lx^2/2} \{2^{n+1}\, (n+1)!\, C_{n+1}^2 + (n+1)^{-3/2}\} = e^{-lx^2/2}\, T_n \end{aligned}$$

where $l < 1$ and K' are suitably chosen positive constants. Since $\sum_1^\infty e^{-lx^2/2} T_n$ is convergent uniformly $-\infty < x < +\infty$, by Theorem IV, the series

$$\sum_0^\infty c_n\, e^{-x^2} H_n(x)$$

converges uniformly to a continuous function $f^*(x)$. Since $\sum_1^\infty e^{-lx^2/2} T_n = Ce^{-lx^2/2}$ where C is a constant, $f^*(x) = O(e^{-lx^2/2})$. Thus we can form the integral $\int_{-\infty}^{+\infty} f^* H_m \, dx$; to evaluate it we write

$$\begin{aligned}
&\frac{1}{2^m m! \sqrt{\pi}} \int_{-\infty}^{+\infty} f^* H_m \, dx - c_m \\
&\qquad = \frac{1}{2^m m! \sqrt{\pi}} \int_{-\infty}^{+\infty} \left(f^* - \sum_0^N c_n e^{-x^2} H_n \right) H_m \, dx \\
&\qquad = \frac{1}{2^m m! \sqrt{\pi}} \int_{-\infty}^{+\infty} H_m \sum_{N+1}^\infty c_n e^{-x^2} H_n \, dx \\
&\qquad = O\left(\int_{-\infty}^{+\infty} e^{-lx^2/2} \cdot |H_m| \cdot \sum_{N+1}^\infty T_n \, dx \right) \\
&\qquad = O\left(\int_{-\infty}^{+\infty} e^{-lx^2/2} |H_m| \, dx \right) \cdot o(1) = o(1)
\end{aligned}$$

as $N \to \infty$. Thus

$$\int_{-\infty}^{+\infty} (f^* - e^{-x^2} f) H_n \, dx = 0, \qquad n = 0, 1, 2, \cdots.$$

By Theorem III the function $\varphi \equiv f^* - e^{-x^2} f$ vanishes identically. The present theorem follows at once.

It is to be noticed that the identity

$$f \equiv e^{x^2} f^* = O(e^{(1-l/2)x^2})$$

where l can be any positive number less than one implies a consequence of the two hypotheses laid on f in the statement of the theorem: any function satisfying them must be $O(e^{kx^2})$ no matter what value greater than one-half is assigned to k.

We turn next to the expansion of a function which has a finite discontinuity at a single point $x = a$. To carry through the discussion we need two lemmas.

LEMMA III. *With the notation* $K_n(x, t) = \frac{1}{2^{n+1} n! \sqrt{\pi}} [H_{n+1}(x) H_n(t) - H_n(x) H_{n+1}(t)]$ *we have*

$$\text{(a)} \quad \frac{1}{\sqrt{\pi}} H_0(x) H_0(t) + \frac{1}{2\sqrt{\pi}} H_1(x) H_1(t) + \cdots + \frac{1}{2^n n! \sqrt{\pi}} H_n(x) H_n(t) = \frac{K_n(x, t)}{x - t};$$

(b) $\int_{-\infty}^{+\infty} e^{-t^2} \frac{K_n(x,t)}{x-t}\,dt = 1;$

(c) $lim_{n\to\infty} e^{-x^2}\int_a^b e^{-t^2} K_n(x,t)\,dt = 0, \quad -\infty < a < b < +\infty,$ *uniformly* $-\infty < x < +\infty;$

(d) $e^{-x^2}\int_a^b e^{-t^2}|K_n(x,t)|\,dt \leqq G,$ *independent of* a, b, n, *and* x, $-\infty < x < +\infty.$

The identity (a) is the analogue of Christoffel's formula in the theory of Legendre polynomials and may be demonstrated in similar fashion. To prove it we write k for $n-1$ in (3) and solve for $xH_k(x) = \frac{1}{2}H_{k+1}(x) + kH_{k-1}(x)$. We see that

$$(x-t)\sum_{k=0}^{k=n} \frac{1}{2^k k!\sqrt{\pi}} H_k(x)\,H_k(t) = \frac{1}{2\sqrt{\pi}}[H_1(x)\,H_0(t) - H_0(x)\,H_1(t)]$$
$$+\sum_{k=1}^{k=n}\left\{\frac{1}{2^{k+1}k!\sqrt{\pi}}[H_{k+1}(x)\,H_k(t) - H_k(x)\,H_{k+1}(t)]\right.$$
$$\left.-\frac{1}{2^k(k-1)!\sqrt{\pi}}[H_k(x)\,H_{k-1}(t) - H_{k-1}(x)\,H_k(t)]\right\} = K_n(x,t).$$

We prove (b) at once by integrating term-wise in (a). To show the truth of (c) we note the formula

$$\int_a^b e^{-t^2} H_k(t)\,dt = \int_a^b (-1)^k \frac{d^k}{dt^k} e^{-t^2}\,dt = (-1)^k \frac{d^{k-1}}{dt^{k-1}} e^{-t^2}\Big|_{t=a}^{t=b}$$
$$= -e^{-t^2} H_{k-1}(t)\Big|_a^b = e^{-a^2} H_{k-1}(a) - e^{-b^2} H_{k-1}(b)$$

whence

$$e^{-x^2}\int_a^b e^{-t^2} K_n(x,t)\,dt = e^{-x^2-a^2}\frac{H_{n+1}(x)\,H_{n-1}(a) - H_n(x)\,H_n(a)}{2^{n+1}n!\sqrt{\pi}}$$
$$- e^{-x^2-b^2}\frac{H_{n+1}(x)\,H_{n-1}(b) - H_n(x)\,H_n(b)}{2^{n+1}n!\sqrt{\pi}}.$$

On the right-hand side of the last identity we can substitute the inequalities of Theorem II, obtaining

$$\left|e^{-x^2}\int_a^b e^{-t^2} K_n(x,t)\,dt\right| \leqq [e^{-(x^2+a^2)l/2} + e^{-(x^2+b^2)l/2}]\frac{K'}{\sqrt{n}} \leqq \frac{2K'}{\sqrt{n}}$$

where $l < 1$ and K' are suitably chosen positive constants. Finally to establish (d) we use Theorem II again:

$$e^{-x^2}\int_a^b e^{-t^2}\,|K_n(x,t)|\,dt \leqq \int_a^b e^{-x^2-t^2}\,\frac{|H_{n+1}(x)H_n(t)|+|H_n(x)H_{n+1}(t)|}{2^{n+1}\,n!\,\sqrt{\pi}}\,dt$$
$$\leqq K'\int_a^b e^{-(x^2+t^2)/2}\,dt < G,$$

where G is independent of a, b, n, and x.

LEMMA IV. *If a is any real number, then the function $\Phi_\mu(x,t)$ defined for $-\infty < x < a$, $a \leqq t < +\infty$, by the equations*

$$\Phi_\mu(x,t) = \frac{1}{x-t_k}, \quad t_{k-1} \leqq t < t_k, \quad k = 1, 2, \cdots, \mu,$$

where $t_0 = a$, $t_\mu = a + \frac{1}{\eta}$, $t_k - t_{k-1} = \frac{t_\mu - t_0}{\mu} = \frac{1}{\mu\eta}$,

$$\Phi_\mu(x,t) = 0, \quad t_u \leqq t < +\infty,$$

satisfies the inequality

$$\left|\Phi_\mu(x,t) - \frac{1}{x-t}\right| < \eta, \quad -\infty < x \leqq a-\varepsilon, \quad a \leqq t < +\infty,$$

for μ sufficiently large.

For $t \geqq t_\mu$, $x \leqq a-\varepsilon$, we have

$$\left|\frac{1}{x-t}\right| \leqq \frac{1}{t_\mu - a + \varepsilon} = \frac{\eta}{1+\eta\varepsilon} < \eta.$$

Then by the law of the mean

$$\frac{1}{x-t} - \frac{1}{x-t_k} = \frac{1}{(x-t')^2}(t-t_k) \qquad t' \text{ on } (t_k, t).$$

Consequently

$$\left|\frac{1}{x-t} - \frac{1}{x-t_k}\right| \leqq \frac{1}{\varepsilon^2}\,\frac{t_\mu - t_0}{\mu} < \eta, \quad t_{k-1} \leqq t \leqq t_k, \quad x \leqq a-\varepsilon,$$

or

$$\left|\frac{1}{x-t} - \Phi_\mu(x,t)\right| < \eta, \quad -\infty < x \leqq a-\varepsilon, \quad a \leqq t < +\infty,$$

for μ sufficiently large.

We are now able to prove Theorem VI.

THEOREM VI. *If $f(x)$ is defined by the equations*

$$f(x) \equiv 0, \quad -\infty < x < a,$$
$$f(a) = \frac{1}{2},$$
$$f(x) \equiv 1, \quad a < x < +\infty,$$

then the series

$$c_0 e^{-x^2} H_0(x) + c_1 e^{-x^2} H_1(x) + \cdots + c_n e^{-x^2} H_n(x) + \cdots,$$
$$c_n = \frac{1}{2^n n! \sqrt{\pi}} \int_a^\infty e^{-x^2} H_n(x)\, dx,$$

converges to $e^{-x^2} f(x)$ everywhere, the convergence being uniform on the ranges $-\infty < x \leqq a - \varepsilon$, $a + \varepsilon \leqq x < +\infty$; and the series

$$c_0 H_0(x) + c_1 H_1(x) + \cdots + c_n H_n(x) + \cdots$$

converges to $f(x)$ everywhere, the convergence being uniform on any closed finite interval for which $x = a$ is an exterior point.

Having fixed upon the quantity $\varepsilon > 0$, and having determined the positive number η so that $G\eta < \frac{\delta}{2}$, where δ is a preassigned positive constant and G is the constant of Lemma III (d), we fix on the integer μ of Lemma IV so that

$$\left| \frac{1}{x-t} - \Phi_\mu(x, t) \right| < \eta, \quad -\infty < x \leqq a - \varepsilon, \quad a \leqq t < +\infty.$$

Now we see that

$$\begin{aligned}
|c_0 e^{-x^2} H_0(x) + \cdots + c_n e^{-x^2} H_n(x)| &= \left| e^{-x^2} \int_a^\infty e^{-t^2} \frac{K_n(x, t)}{x-t}\, dt \right| \\
&\leqq \left| e^{-x^2} \int_a^\infty e^{-t^2} \left[\frac{1}{x-t} - \Phi_\mu(x, t) \right] K_n(x, t)\, dt \right| \\
&\quad + \left| e^{-x^2} \int_a^{t_\mu} e^{-t^2} \Phi_\mu(x, t) K_n(x, t)\, dt \right| \\
&\leqq G\eta + \left| \sum_{k=1}^{\mu} \frac{e^{-x^2}}{x - t_k} \int_{t_{k-1}}^{t_k} e^{-t^2} K_n(x, t)\, dt \right| \\
&\leqq \frac{\delta}{2} + \frac{1}{\varepsilon} \left| \sum_{k=1}^{\mu} e^{-x^2} \int_{t_{k-1}}^{t_k} e^{-t^2} K_n(x, t)\, dt \right| < \delta, \quad n > N,
\end{aligned}$$

independent of x, $-\infty < x \leqq a - \varepsilon$, by Lemma III (c). Next, for x on the interval $a + \varepsilon \leqq x < +\infty$, we write

$$\begin{aligned}
c_0 e^{-x^2} H_0(x) + \cdots + c_n e^{-x^2} H_n(x) &= e^{-x^2} \int_a^\infty e^{-t^2} \frac{K_n(x, t)}{x-t}\, dt \\
&= e^{-x^2} - e^{-x^2} \int_{-\infty}^a e^{-t^2} \frac{K_n(x, t)}{x-t}\, dt
\end{aligned}$$

with the help of Lemma III (b). Thus, by (7),

$$c_0 e^{-x^2} H_0(x) + \cdots + c_n e^{-x^2} H_n(x) - e^{-x^2} = -e^{-x^2} \int_{-\infty}^{+a} e^{-t^2} \frac{K_n(x, t)}{x - t}\, dt$$
$$= -e^{-x'^2} \int_{a'}^{\infty} e^{-t'^2} \frac{K_n(x', t')}{x' - t'}\, dt'$$

where $x' = -x$, $t' = -t$, $a' = -a$. The reasoning applied above can be repeated for the integral on the right; we see that

$$|c_0 e^{-x^2} H_0(x) + \cdots + c_n e^{-x^2} H_n(x) - e^{-x^2}| < \delta, \qquad n > N',$$

independent of x, $a + \varepsilon \leqq x < +\infty$. Consequently the series converges to $e^{-x^2} f(x)$ uniformly, $-\infty < x \leqq a - \varepsilon$, $a + \varepsilon \leqq x < +\infty$. The corresponding statements can then be made about the second series.

It remains for us to establish the convergence of the two series when $x = a$. To discuss this question we write

$$S_n(a) = c_0 H_0(a) + \cdots + c_n H_n(a)$$

and show that $S_n(a) \to \frac{1}{2}$ when $n \to \infty$. We have by (4)

$$c_0 = \frac{1}{\sqrt{\pi}} \int_a^{\infty} e^{-x^2}\, dx,$$

$$c_n = \frac{1}{2^n n! \sqrt{\pi}} \int_a^{\infty} e^{-x^2} H_n(x)\, dx = \frac{e^{-a^2} H_{n-1}(a)}{2^n n! \sqrt{\pi}},$$

for $n = 1, 2, 3, \cdots$. Thus

$$S_n(a) = \frac{1}{\sqrt{\pi}} \int_a^{\infty} e^{-x^2}\, dx + \sum_1^n \frac{e^{-a^2} H_k(a) H_{k-1}(a)}{2^k k! \sqrt{\pi}},$$

$$S_n(0) = \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-x^2}\, dx = \frac{1}{2}.$$

Differentiating with respect to a we see from (1) that

$$S_n' = -\frac{1}{\sqrt{\pi}} a e^{-a^2} H_1(a)$$
$$+ \sum_2^n \frac{e^{-a^2} [-2a H_k(a) H_{k-1}(a) + 2(k-1) H_{k-2}(a) H_k(a)]}{2^k k! \sqrt{\pi}}$$
$$+ \sum_2^n \frac{e^{-a^2} H_{k-1}^2(a)}{2^{k-1} (k-1)! \sqrt{\pi}}.$$

On rearranging the indices in the last sum, we find

$$S_n' = \sum_2^n \frac{e^{-a^2} H_k(a)\,[H_k(a) - 2a\,H_{k-1}(a) + 2(k-1)\,H_{k-2}(a)]}{2^k\,k!\,\sqrt{\pi}} - \frac{e^{-a^2} H_n^2(a)}{2^n\,n!\,\sqrt{\pi}}.$$

By virtue of (3) this equation reduces to

$$S_n' = \frac{-e^{-a^2} H_n^2(a)}{2^n\,n!\,\sqrt{\pi}}$$

whence

$$S_n(a) = \frac{1}{2} - \frac{1}{2^n\,n!\,\sqrt{\pi}} \int_0^a e^{-x^2} H_n^2(x)\,dx.$$

For fixed a, therefore,

$$S_n(a) - \frac{1}{2} = O\left(\frac{1}{\sqrt{n}}\right)$$

in accord with the inequalities of Theorem II. This completes the proof of the theorem.

Finally we can discuss the expansions for any function representable as a linear combination of a finite set of functions of the types described in Theorems V and VI. We thus have a fairly comprehensive treatment of series of Hermite polynomials for functions of this simple type.

COLUMBIA UNIVERSITY,
NEW YORK CITY.

A GENERALIZATION OF THE CALCULUS OF FINITE DIFFERENCES TO INCLUDE THE DIFFERENTIAL CALCULUS.*

By J. P. Ballantine.

As usually treated, the calculus of finite differences includes the differential calculus only as a limiting case. A difference quotient is defined for any set of values which are distinct, and when the theory is carried further to include the case of coincident values, the limit of the difference quotient is considered.† It is the purpose of the present paper to show how the difference quotient may be introduced so that its definition will hold for any set of values $x_1, x_2, \cdots, x_{n+1}$, whether these are distinct or not. For the purpose of notation we assume that x_i appears k_i times, where i takes on values from 1 to h, and $\sum k_i = n+1$.

It is not absolutely necessary for our treatment, but will obviate the necessity of certain clumsy reasoning if we assume the following well known theorem:

If $x_1, x_2, \cdots, x_h$ are any distinct values and $k_1, k_2, \cdots, k_h$ are any positive or zero integers whose sum is $n+1$, and if $f(x_1), f'(x_1), \cdots, f^{(k_1-1)}(x_1), f(x_2), \cdots, f^{(k_h-1)}(x_h)$ are any real numbers, then there exists one and only one polynomial $f(x)$ of degree n or less which together with its derivatives takes on the prescribed values. For convenience we shall denote this polynomial by the symbol

$$\left(\frac{f(x_1),\ f'(x_1),\ \cdots,\ f^{(k_1-1)}(x_1),\ f(x_2),\ \cdots,\ f^{(k_h-1)}(x_h)}{x_1\ ,\quad x_1\ ,\cdots,\quad x_1\quad ,\quad x_2\ ,\cdots,\quad x_h}\right).$$

The set of values occurring in the lower row of the above symbol, referred to as the denominator of the symbol, may be denoted by a single letter σ. The *degree* of any set σ is one less than the number of elements, namely n. The *evanescent* of any set σ, denoted by $\mathfrak{E}_\sigma(x)$, is defined as the product of the factors formed by subtracting each element of σ from x, namely

$$\mathfrak{E}_\sigma(x) = (x-x_1)^{k_1}\ (x-x_2)^{k_2} \cdots (x-x_h)^{k_h}.$$

The degree of the evanescent is, of course, one greater than the degree of the set.

* Presented to the American Mathematical Society, December 28, 1923. Received May 28, 1927.

† H. L. Smith, Annals of Math., vol. 25 (1923) p. 123.

A function $f(x)$ is said to be defined at a set σ if the function is defined and its first $k_i - 1$ derivatives exist for the values x_i of the set.

If $f(x)$ is any function defined at a set σ, then the *secant* to $f(x)$ at σ, denoted by $f_\sigma(x)$, is defined as the above uniquely existing polynomial. The *difference quotient* of $f(x)$ at σ, written $f(\sigma)$, is defined as the nth derivative of the secant $f_\sigma(x)$.

It is obvious, but logically necessary to point out, that in the case when $n = 0$, and σ consists of a single element x_1 appearing once, $f(\sigma)$ and $f(x_1)$ are numerically equal, and hence no confusion need arise by replacing σ by its element(s), a procedure which is always followed when σ is a particular numerical set with a small number of elements. Also, in case $n = 0$, we consider that the nth derivative of a function is the function itself, so that $f_\sigma(x) = f(\sigma)$.

THEOREM 1. *If a function $f(x)$ is defined at a set σ of degree n, and σ_1 is the set obtained by striking x_1 once out of σ, then*

$$f_\sigma(x) = f_{\sigma_1}(x) + \frac{f(\sigma)\ \mathfrak{E}_{\sigma_1}(x)}{n!}.$$

Since $f_\sigma(x)$ and $f_{\sigma_1}(x)$ have the same values for the values of σ_1,

$$f_\sigma(x) - f_{\sigma_1}(x) = A\,\mathfrak{E}_{\sigma_1}(x).$$

Since the degree of A is at most 0, A is a constant.

Differentiating n times, since $f_{\sigma_1}(x)$ is a polynomial of degree at most $n-1$, and $\mathfrak{E}_{\sigma_1}(x)$ is of degree at most n and its highest term has a coefficient 1, we have by the definition of difference quotient

$$\frac{d^n}{dx^n} f_\sigma(x) = f(\sigma) = A \cdot n!$$

Eliminating A from the last two equations, the theorem is proved.

THEOREM 2. *If $f(x)$ is a function defined for a set σ of degree n, and σ_1 and σ_2 are the sets obtained from σ by striking out once x_1 and x_2, respectively, where $x_1 \neq x_2$, then*

$$f(\sigma) = \frac{f(\sigma_1) - f(\sigma_2)}{\left(\frac{x_2 - x_1}{n}\right)}.$$

From Theorem 1,

$$f_\sigma(x) = f_{\sigma_1}(x) + f(\sigma)\ \mathfrak{E}_{\sigma_1}(x),$$
$$f_\sigma(x) = f_{\sigma_2}(x) + f(\sigma)\ \mathfrak{E}_{\sigma_2}(x),$$

whence

$$f_{\sigma_1}(x) - f_{\sigma_2}(x) = \frac{f(\sigma)}{n!}\,(\mathfrak{E}_{\sigma_2}(x) - \mathfrak{E}_{\sigma_1}(x)).$$

If σ_{12} denotes the set obtained from σ by striking out once both x_1 and x_2, then

$$\mathfrak{E}_{\sigma_1}(x) = (x - x_2)\ \mathfrak{E}_{\sigma_{12}}(x),$$
$$\mathfrak{E}_{\sigma_2}(x) = (x - x_1)\ \mathfrak{E}_{\sigma_{12}}(x).$$

Hence

$$f_{\sigma_1}(x) - f_{\sigma_2}(x) = \frac{1}{n!}\, f(\sigma)\ \mathfrak{E}_{\sigma_{12}}(x) \cdot (x_2 - x_1).$$

Differentiating $n-1$ times, we have

$$f(\sigma_1) - f(\sigma_2) = \frac{f(\sigma)\ \{(n-1)!\}\ (x_2 - x_1)}{n!},$$

and obtain the theorem by solving for $f(\sigma)$.

THEOREM 3. *If $f(x)$ is defined at a set σ consisting of x_1 appearing $n+1$ times, then*

$$f(\sigma) = f^{(n)}(x).$$

The proof follows directly by applying Taylor's Theorem.

The theory is now complete. In any set σ of degree $n \geqq 0$, either two of its elements are unequal and Theorem 2 applies, or else Theorem 3 applies. Hence every difference quotient can be expressed in terms of difference quotients of lower order and ultimately in terms of the values of the function and its derivatives. In a similar way Theorem 1 expresses the secant at any set σ in terms of lower degree secants. Theorem 1 contains as special cases Newton's forward and Newton's backward formulas, and also Gauss' forward and backward formulas, and Taylor's expansion.

UNIVERSITY OF WASHINGTON, SEATTLE.

RELATIONS SATISFIED BY COEFFICIENTS OF PERIODIC SOLUTIONS.*

By W. J. Trjitzinsky.

A general procedure will be indicated applicable to an equation of nth order with periodic coefficients containing parameters, by means of which relations satisfied by the coefficients of a periodic solution are derivable

In the case of the Mathieu differential equation

$$\frac{d^2y}{dx^2} + (a + k^2\cos^2 x)\cdot y = 0 \tag{1}$$

which is of importance in several physical problems, there are infinitely many determinations of a in terms of k^2 each in a form of a power series in k^2; for each of these determinations the equation possesses a periodic solution (Mathieu function). It was shown by McDonald† that the coefficients of such a solution, when arranged as a trigonometrical series, can be determined without actually knowing the expression of a in terms of k^2. Whittaker‡ has given a different method; but by means of it only a few solutions are obtainable in this way. With the assumption of these facts, I have considered§ any one of the periodic solutions of the form:

$$y_1 = b_0 + b_1\cos 2x + b_2\cos 4x + \cdots + b_n\cos 2nx + \cdots$$

where b_i are analytic functions of k^2; denoting $\partial y_1/\partial(k^2) = w_1$ an equation was obtained

$$\frac{d^2w_1}{dx^2} + (a + k^2\cos^2 x)\cdot w_1 = -\left(\frac{da}{d(k^2)} + \cos^2 x\right) y_1;$$

solving it by the method of variation of parameters and noting that w_1 is periodic, the relation was derived

$$\left(\frac{da}{d(k^2)} + \frac{1}{2}\right)\cdot[2b_0^2 + b_1^2 + b_2^2 + \cdots] + \frac{1}{2}[2b_0b_1 + b_1b_2 + \cdots] = 0 \tag{2}$$

which is characteristic of Mathieu functions, and by means of it certain propositions were established of theoretical interest.

* Received May 3, 1927.

† J. H. McDonald. Manuscript.

‡ E. T. Whittaker. Proc. Edinb. Math. Soc., 32 (1914), pp. 75–80.

§ W. J. Trjitzinsky. The elliptic cylinder differential equation. Thesis (1926). Library University of California.

Let us consider a differential equation:

$$p_0 y^{(n)} + p_1 y^{(n-1)} + \cdots + p_n y = 0 \tag{3}$$

where p_i are convergent trigonometrical series of period 2π containing parameters, k being one of them. Let us suppose that under certain restrictions upon the parameters there is a periodic solution

$$y_1 = b_0 + \sum_{i=1}^{\infty} (b_i \cos ix + c_i \sin ix);$$

letting $\partial y_1/\partial k = w_1$ and differentiating (3) with respect to k, we have

$$p_0 w_1^{(n)} + p_1 w_1^{(n-1)} + \cdots + p_n w_1 = -\left[\frac{\partial p_0}{\partial k} y_1^{(n)} + \cdots + \frac{\partial p_n}{\partial k} y_1\right] = f(x). \tag{4}$$

The natural extension of the method used in deriving (2) would be to solve (4) by the method of variation of parameters and introduce the condition that w_1 is periodic. In the general case this, however, would be difficult to perform on account of the complexity of the expressions. The results of Fite* can be used with advantage. Suppose that (3) has a periodic solution y_1, then, by his Theorem I the adjoint equation

$$(p_0 z)^{(n)} - (p_1 z)^{(n-1)} + \cdots + (-1)^n (p_n z) = 0 \tag{5}$$

has periodic solutions $z_1, z_2, \cdots, z_i$ $(i \geqq 1)$; by his Theorem II the necessary and sufficient condition that w_1 as determined by (4) should be periodic is that

$$\int_x^{x+2\pi} z_j f(x)\, dx = 0, \qquad 1 \leqq j \leqq i \tag{6}$$

for some z_j, provided that not all solutions of (3) are periodic (which will be supposed to be the case).

Then is at least one equation of type (6), say

$$\int_x^{x+2\pi} z_k f(x)\, dx = 0. \tag{6_1}$$

Under our suppositions $[z_k f(x)]$ is a trigonometric series, and (6_1) amounts to the statement that the constant term in that series is zero; thus, if

$$z_k = e_0 + \sum_{i=1}^{\infty} (e_i \cos ix + d_i \sin ix)$$

* W. B. Fite. Periodic solutions of linear differential equations, Annals of Math., (2), vol. 28 (1926), pp. 59–64.

(6_1) is found to be equivalent to

$$\text{(A)} \qquad \begin{aligned} &k_0 e_0 + k_1 e_1 + k_2 e_2 + \cdots + k_n e_n + \cdots, \\ &r_1 d_1 + r_2 d_2 + \cdots + r_n d_n + \cdots = 0 \end{aligned}$$

where

$$k_i = \frac{1}{\pi}\int_{-\pi}^{\pi}\left[\frac{\partial p_0}{\partial k}(y_1)^{(n)} + \cdots + \frac{\partial p_n}{\partial k} y_1\right]\cdot \cos ix\, dx$$

$$r_i = \frac{1}{\pi}\int_{-\pi}^{\pi}\left[\frac{\partial p_0}{\partial k}(y_1)^{(n)} + \cdots + \frac{\partial p_n}{\partial k} y_1\right]\cdot \sin ix\, dx.$$

(A) is a relation satisfied by the coefficients b_i, c_i of y_1—and e_i, d_i of z_k. The quantities b_i, c_i, e_i, d_i are implicitely functions of the parameters, and (A) is bilinear in them, being besides a differential equation in the parameter k.

Applying (A) to the Mathieu differential equation (differentiating with respect to k^2) we observe that only one of its solutions can be periodic,* that it is self-adjoint, and thus we obtain the relation (2), which was obtained previously by the method of variation of parameters.

Let us apply (A) to an equation

$$\text{(7)} \qquad \frac{d^2y}{dx^2} + [\theta_0(k) + \theta_1(k)\cos 2x + \theta_2(k)\cos 4x + \cdots + \theta_n(k)\cos 2nx + \cdots] \times y = 0;$$

where θ_i are analytic functions of a parameter k and may contain other parameters, $\sum_{i=0}^{\infty} |\theta_i|$ and $\sum_{i=0}^{\infty} |\partial\theta_i/\partial k|$ converge, and $\theta_i(k)$ are such functions that when (7) has a periodic solution, say

$$y_1 = b_0 + b_1\cos 2x + b_2\cos 4x + \cdots + b_n\cos 2nx + \cdots,$$

its second solution is non periodic (this happens, for example, when (7) is a Mathieu equation). When the parameters are absent (7) is Hill's equation of Lunar theory. The equation is self adjoint, so $z_k = z_1 = y_1$ and (A) becomes:

$$\begin{aligned} &\frac{\partial\theta_0}{\partial k}[2b_0^2 + b_1^2 + b_2^2 + \cdots] + \frac{\partial\theta_1}{\partial k}[2b_0b_1 + b_1b_2 + b_2b_3 + b_3b_4 + \cdots] \\ &+ \frac{\partial\theta_2}{\partial k}\left[\frac{b_1^2}{2} + 2b_0b_2 + b_1b_3 + \cdots\right] + \frac{\partial\theta_3}{\partial k}[2b_0b_3 + b_1b_2 + b_1b_4 + b_2b_5 + b_3b_6 + \cdots] \\ &+ \frac{\partial\theta_4}{\partial k}\left[\frac{b_2^2}{2} + 2b_0b_4 + b_1b_3 + b_1b_5 + b_2b_6 + b_3b_7 + \cdots\right] + \cdots = 0; \end{aligned} \qquad \text{(8)}$$

* J. H. Mc Donald. Manuscript.

in (8) b_i are unknown functions of the parameters; (in some cases, as in Mathieu equation, they may be known without knowing the restriction on the parameters producing periodic solutions).

When θ_i $(i \geqq 2)$ are independent of k, b_i satisfy a relation similar to (2), namely

$$(9)\quad \frac{\partial \theta_0}{\partial k}[2b_0^2+b_1^2+b_2^2+\cdots]+\frac{\partial \theta_1}{\partial k}[2b_0b_1+b_1b_2+b_2b_3+b_3b_4+\cdots]=0.$$

In (8) the multipliers of $\partial\theta_i/\partial k$ $(i=0,1,2,\cdots)$ are coefficients of $\cos 2ix$ in $\left[\sum_{j=0}^{\infty} b_j \cos 2jx\right]^2$, and are convergent for the following reasons.

In Whittaker's discussion of Hill's equation* θ_i are constants, but the results which we are about to quote still hold, when θ_i depend on a parameter k. He states that the solution of the equation can be given in the form $u = e^{\mu x}\cdot\sum_{-\infty}^{+\infty} c_n e^{2nix}$. When it is periodic, we have,† say, $\sum_{-\infty}^{+\infty} c_n e^{2nix}$. In either case $\sum_{n=-\infty}^{+\infty} |n^2 c_n|$ converges. The solution according to our hypothesis, is included in the form $\sum_{-\infty}^{+\infty} c_n e^{2nix}$ when $c_n = c_{-n} = b_n$ $(n = 1, 2, \cdots)$, $c_0/2 = b_0$; then $\sum_{-\infty}^{+\infty} |n^2 c_n| = 2\sum_{n=0}^{\infty} |n^2 b_n|$, and, consequently, $\sum_{n=0}^{\infty} |n^2 b_n|$ and $\sum_{n=0}^{\infty} |b_n|$ converge.

In (8) the multipliers of $\partial\theta_i/\partial k$ are series each term of which is a product of two factors. Taking any multiplier, in each of its products a factor can be picked out so that they will form a sequence or subsequence of $b_0, b_1, b_2, \cdots$ (where some terms may be repeated a finite number of times). The sum of these factors converges, since $\sum_{n=0}^{\infty} |b_n|$ does; the remaining factors being bounded, the convergence of all the multipliers of $\partial\theta_i/\partial k$ $(i = 0, 1, 2, \cdots)$ will be secured, and since $\sum_{i=0}^{\infty} |\partial\theta_i/\partial k|$ converges, the convergence of the left member of (8) will follow.

* Whittaker and Watson. Modern analysis, p. 414.

† Generally, a periodic solution will be $u = \sum_{n=-\infty}^{+\infty} c_n e^{2(n+k)ix}$, k being an integer $(\mu = 2ki)$. The discussion would be essentially the same, when $k \neq 0$, convergence of $\sum_{n=0}^{\infty} |b_n|$ being secured by that of $\sum_{n=0}^{\infty} |b_n|(k^2+n^2)$.

AN INTEGRAL EQUATION WITH AN ASSOCIATED INTEGRAL CONDITION.*

By Laura Guggenbühl.

Introduction. Several articles have appeared which are concerned with the solution of an integral equation and an auxiliary integral condition, or with the solution of a differential equation with associated integral conditions instead of the usual point boundary conditions. Bateman,† Cairns,‡ and Pell§ have treated the former, and von Mises‖ the latter.

In the present paper we consider a natural extension of the problem considered by Pell. Thus, we consider the integral equation

$$u(s) = \lambda \int_a^b K(sx)\, u(x)\, dx + \mu p(s) \tag{1}$$

with the auxiliary condition

$$\int_a^b Au = 0, \tag{2}$$

where $K(sx)$ is a continuous symmetric kernel on the square $a \leqq x \leqq b$, $a \leqq s \leqq b$, $p(s)$ and $A(s)$ are two continuous functions on an interval $a \leqq s \leqq b$, $\int_a^b AK \not\equiv 0$, and λ, μ are two arbitrary parameters. First we develop a sufficient condition for the existence of real solutions of (1) and (2). In Section 3 we discuss the usual expansion theorems using the theory of symmetrizing matrices developed in Section 2.

Section 1. General Discussion of Problem.

Reduction to a single integral equation. In the first place, we show that (1) and (2) are equivalent to a certain integral equation without auxiliary condition if $\int_a^b Ap \neq 0$.

Multiplying (1) by $A(s)$ and integrating,¶ we obtain

$$\iint AKu + \mu \int Ap = 0 \tag{3}$$

for any solution of (1) and (2).

---

* Received June 11, 1927.

† Bateman, Trans. Cambr. Phil. Soc., vol. 21 (1908), pp. 123–128.

‡ Cairns, "Die Anwendung der Integralgleichungen auf die zweite Variation bei isoperimetrischen Problemen", Diss., Göttingen, 1907.

§ Pell, Bull. Amer. Math. Soc., (2), vol. 16 (1910), pp. 412–415.

‖ von Mises, "Beitrag zum Oszillationsproblem", Heinrich Weber-Festschrift, 1912.

¶ The limits of integration will be omitted when there is no ambiguity.

It follows that

$$\mu = -\lambda \frac{\iint AKu}{\int Ap}. \tag{3'}$$

Applying substitution (3') to (1) we obtain

$$u(s) = \lambda \int \left[K(sx) - \frac{p(s) \int A(t)\, K(tx)\, dt}{\int Ap} \right] u(x)\, dx. \tag{4}$$

Any solution of (1) and (2) for a certain λ and a corresponding μ satisfies (4) for this λ. Conversely, any solution of (4) for a certain λ satisfies (1) and (2) for this λ and a corresponding μ. The original problem, then, is equivalent to a discussion of integral equation (4), if $\int Ap \neq 0$.*

Three special cases. Some cases require no new theory. For example, the function

$$p(s) = \int A(t)\, K(ts)\, dt$$

leads immediately to the symmetric continuous kernel

$$K(sx) - \frac{\int A(t)\, K(ts)\, dt \int A(r)\, K(rx)\, dr}{\iint AKA}.$$

If K is a positive definite kernel, the particular case

$$p(s) = \int A(t)\, K^{(2)}(ts)\, dt$$

leads to the symmetrizable kernel

$$K(sx) - \frac{\int A(t)\, K^{(2)}(ts)\, dt \int A(r)\, K(rx)\, dr}{\iint AK^{(2)}A}.$$

with K as symmetrizing kernel.

---

* It can easily be seen that the following equation and auxiliary condition form an adjoint of (1) and (2),

$$v(x) = \lambda \int v(s)\, K(sx)\, ds + \mu \int A(t)\, K(tx)\, dt \quad \text{and} \quad \int rv = 0,$$

if $p(s) = \int r(t)\, K(ts)\, dt$, for in this case, we can show that these two equations are equivalent to the single integral equation

$$v(x) = \lambda \int v(s) \left[K(sx) - \frac{p(s) \int A(t)\, K(tx)\, dt}{\int Ap} \right] ds.$$

However, in what follows we shall not need to impose the condition $p = \int rK$.

Also, if the kernel K is positive definite, and if

$$p(s) = A(s)$$

we obtain the symmetrizable kernel

$$K(sx) - \frac{A(s)\int A(t)\,K(tx)\,dt}{\int A^2},$$

with K as symmetrizing kernel.*

All of these cases come under the later discussion, but in each either the kernel itself is symmetric or a symmetrizing *kernel* exists, and hence we can state for these special cases more general expansion theorems than those given in Section 3.

Preliminary discussion of characteristic numbers for $p_i \neq 0$. In what follows, *we shall assume that* $\int Ap = 1$. We may then write (4) as

$$u(s) = \lambda \int \left[K(sx) - p(s)\int A(t)\,K(tx)\,dt\right] u(x)\,dx. \tag{5}$$

Let φ_i be a characteristic function of $K(sx)$ corresponding to the characteristic number $\varkappa_i$. If there exists a solution $u(s)$ of (5) for $\lambda = \lambda_\alpha$, then

$$\int u_\alpha \varphi_i = \lambda_\alpha \int\int \varphi(s)\,K(sx)\,u_\alpha(x)\,dx\,ds - \lambda_\alpha \int p\varphi_i \int\int AKu_\alpha$$

or

$$u_{\alpha i} = \lambda_\alpha \frac{u_{\alpha i}}{\varkappa_i} - \lambda_\alpha p_i \sum_i \frac{a_i u_{\alpha i}}{\varkappa_i}$$

where

$$u_{\alpha i} = \int u_\alpha \varphi_i, \quad a_i = \int A\varphi_i, \quad p_i = \int p\varphi_i.$$

It follows that

$$u_{\alpha i}\left[\frac{\varkappa_i - \lambda_\alpha}{\varkappa_i \lambda_\alpha}\right] = p_i M_\alpha, \qquad M_\alpha = \sum_i \frac{a_i u_{\alpha i}}{\varkappa_i}.$$

From (3) we see that M_α can be zero only for $\int Ap = 0$ or for $\mu = 0$. We have already excluded the former possibility. If $\mu = 0$ we see from (1) that our problem is reduced to the consideration of the equation

$$u = \lambda \int Ku,$$

* Pell, l. c., For $p(s) = A(s)$ equations (1) and (2) are reduced to the symmetric case even if K is not positive definite.

i. e. the ordinary homogeneous integral equation with symmetric kernel. Therefore the case $\mu = 0$ requires no further investigation. Then, if we assume for the present

$$p_i \neq 0,$$

we may write, when $\varkappa_i - \lambda_\alpha \neq 0$

$$u_{\alpha i} = \frac{p_i \varkappa_i \lambda_\alpha}{\varkappa_i - \lambda_\alpha} M_\alpha.$$

If we make the additional assumption that $p(s) = \sum_i p_i \varphi_i(s)$, and that the series converges uniformly, we see from (5) that the series

$$u_\alpha(s) = \sum_i u_{\alpha i} \varphi_i(s)$$

is uniformly convergent. That is, if a solution $u_\alpha(s)$ of (5) exists for $\lambda = \lambda_\alpha$, it must have the form

$$u_\alpha(s) = \lambda_\alpha M_\alpha \sum_i \frac{p_i \varkappa_i \varphi_i(s)}{\varkappa_i - \lambda_\alpha}.$$

Since $\int A u_\alpha = 0$ for any solution of (5)

$$\sum_i \frac{a_i p_i \varkappa_i}{\varkappa_i - \lambda_\alpha} = 0 \tag{6}$$

for any value λ_α of λ for which there exists a solution of (5). If we cannot say that the series $\sum_i p_i \varphi_i(s)$ converges uniformly, but if we do know that $p(s)$ is determined by its Fourier coefficients with respect to $\varphi_i(s)$, i. e. that $\int p \varphi_i^* = 0$, where $\varphi_i(s)$ and $\varphi_i^*(s)$ form a closed system, we can modify the above discussion as follows. Since φ_i and φ_i^* form a closed system,

$$\int A u_\alpha = \sum_i \int A \varphi_i \int u_\alpha \varphi_i + \sum_i \int A \varphi_i^* \int u_\alpha \varphi_i^*.$$

But from (5) it follows that $\int u_\alpha \varphi_i^* = 0$, and therefore we obtain (6) again. Conversely, the function

$$u_\alpha(s) = c \lambda_\alpha \sum_i \frac{p_i \varkappa_i \varphi_i(s)}{\varkappa_i - \lambda_\alpha}, \tag{7}$$

if the series on the right converges uniformly (which is so if $\sum_i p_i \varphi_i(s)$ converges uniformly), is a solution of (5) when λ_α is a value of λ for which

$$(6') \qquad D(\lambda) = \sum_i \frac{a_i p_i \varkappa_i}{\varkappa_i - \lambda} = 0.$$

This shows, then, that the roots of the Fredholm Determinant for the kernel

$$L(sx) = K(sx) - p(s) \int A(t)\, K(tx)\, dt,$$

which are not characteristic numbers of K, are also roots of the equation $D(\lambda) = 0$, and conversely.

A graphic representation of this function shows immediately that there is a real characteristic number of (5) whenever $a_i p_i \varkappa_i$ and $a_{i+1} p_{i+1} \varkappa_{i+1}$ have the same sign. Analytically we prove the existence of these characteristic numbers by changes in sign of $D(\lambda)$. If some of the characteristic numbers of K are of more than single multiplicity of the Fredholm Determinant for K, we combine the terms of $D(\lambda)$ which have the same denominator, and make the indicated change in the statement of the above sufficient condition. The graphic representation also suggests that if

$$(8) \qquad a_i p_i \varkappa_i \geqq 0$$

(because of the conditions already imposed not all the $a_i p_i \varkappa_i$ can be zero), the solutions of (6′) and therefore all of the characteristic numbers of (5) are real. The proof follows readily.

For, suppose that $\lambda = c + id$ is a solution of $D(\lambda) = 0$, i. e.

$$\sum_i \frac{a_i p_i \varkappa_i}{\varkappa_i - c - id} = 0.$$

By rationalizing the denominator and separating the result into real and imaginary parts, we see that

$$(9) \qquad \sum_i \frac{a_i p_i \varkappa_i}{(\varkappa_i - c)^2 + d^2} = 0.$$

But (9) is impossible, which means that $D(\lambda)$ can have no imaginary roots under condition (8), and therefore, that all of the characteristic numbers of (5) are real.

THEOREM 1. *If* $\int Ap \neq 0$, *the function* $p(s)$ *is determined by its Fourier coefficients with respect to* $\varphi_i(s)$, *if further* $p_i \neq 0$, $a_i p_i \varkappa_i \geqq 0$, *and the series in* (7) *converges uniformly, then equation* (5) *has real and only real characteristic numbers, and these characteristic numbers separate the characteristic numbers of* K *either individually or in groups.*

Since we now have a condition under which all of the characteristic numbers of (5) are real, the question of a symmetrizing kernel for L naturally arises. If K is closed, and $a_i p_i \varkappa_i \geqq 0$ (from now on we shall assume that not all of the $a_i p_i \varkappa_i$ are zero), and if either the series

$$T(sy) = \sum_i \frac{a_i}{p_i \varkappa_i} \varphi_i(s) \varphi_i(y)$$

or the series

$$\mathcal{L}(sy) = \sum_i \frac{p_i \varkappa_i}{a_i} \varphi_i(s) \varphi_i(y)$$

converges uniformly, we have a desired symmetrizing kernel.

However, we cannot in general say that either of these series converges. For this reason we shall develop a theory of a symmetrizing *matrix*, which is not necessarily bounded, and which has only a point spectrum.

Section 2. Symmetrizing Matrices with Application to Integral Equations.

Symmetrizing matrizes. Suppose that we have a matrix $(l_{\alpha\beta})$ such that $\sum_{\alpha,\beta} l_{\alpha\beta}^2$ converges, and a symmetric matrix $(t_{\alpha\beta})$ which has the following properties:

$$(10)\quad \begin{cases} \text{(a) } \sum_{\alpha,\beta} x_\alpha t_{\alpha\beta} x_\beta \geqq 0 \\ \quad \text{for } \{x_\alpha\} \text{ of finite norm and such that this sum exists, equal to zero only for } x_\alpha \equiv 0, \text{ or for } \{x_\alpha^*\}, \text{ of finite norm and such that } \sum_\beta t_{\alpha\beta} x_\beta^* = 0; \\ \text{(b) } \sum_\beta t_{\alpha\beta}^2 \text{ converges;} \\ \text{(c) } \sum_\nu t_{\alpha\nu} l_{\nu\beta} = \sum_\nu t_{\beta\nu} l_{\nu\alpha} \not\equiv 0, \text{ and of finite norm;} \\ \text{(d) } \sum_\alpha x_\alpha \sum_\beta t_{\alpha\beta} y_\beta = \sum_\beta (\sum_\alpha x_\alpha t_{\alpha\beta}) y_\beta \\ \quad \text{for } \{x_\alpha\}\ \{y_\beta\} \text{ of finite norm, and such that } \sum_\alpha t_{\alpha\beta} x_\alpha \text{ and } \sum_\beta t_{\alpha\beta} y_\beta \text{ are of finite norm:} \end{cases}$$

we shall call the matrix $(t_{\alpha\beta})$ a symmetrizing matrix of the matrix $(l_{\alpha\beta})$.

Properties (10, b) and (10, d) make the matrix $(t_{\alpha\beta})$ of Carlemann's Class 1.† From (10, b) it follows that there exists a spectrum for the matrix $(t_{\alpha\beta})$, in general both a point and a continuous spectrum. However, *in this paper we are concerned with the case in which the matrix* $(t_{\alpha\beta})$ *has only a point spectrum.* On account of (10, a) the characteristic numbers of $(t_{\alpha\beta})$ are positive. We may therefore write

† T. Carlemann, "Sur les Équations Intégrales Singulières a Noyau Réel et Symétrique", 1923.

(11) $$x_{\alpha i} = \tau_\alpha^2 \sum_\nu t_{i\nu} x_{\alpha\nu}$$

where $x_{\alpha i}$ is a system of finite norm with respect to both α and i, and

$$\sum_i x_{\alpha i} x_{\beta i} = e_{\alpha\beta},$$

and

(12) $$\sum_\alpha x_{\alpha i} x_{\alpha j} = e_{ij} - \sum_\alpha x^*_{\alpha i} x^*_{\alpha j},$$

$x^*_{\alpha i}$ an orthogonal system. Build up

(13) $$c_{\alpha\beta} = \sum_{m,n} \frac{x_{\alpha m}}{\tau_\alpha} l_{mn} x_{\beta n} \tau_\beta.$$

From (11) we have

$$c_{\alpha\beta} = \tau_\alpha \sum_{m,n,\nu} t_{m\nu} x_{\alpha\nu} l_{mn} x_{\beta n} \tau_\beta,$$

and from (10, c) and (10, d)

$$\begin{aligned} c_{\alpha\beta} &= \sum_{\nu,n} \tau_\beta x_{\beta\nu} \sum_m t_{m\nu} l_{mn} x_{\alpha n} \tau_\alpha \\ &= c_{\beta\alpha}. \end{aligned}$$

We note next that the supplementary matrix $(c_{\alpha\beta})$ is completely continuous. For

$$\begin{aligned} \sum_{\alpha,\beta=1}^{r} c_{\alpha\beta}^2 &= \sum_{\alpha,\beta=1}^{r} c_{\alpha\beta} c_{\beta\alpha} \\ &= \sum_{\alpha,\beta=1}^{r} \left[\sum_{m,n} \frac{x_{\alpha m}}{\tau_\alpha} l_{mn} x_{\beta n} \tau_\beta\right]\left[\sum_{m,n} \frac{x_{\beta n}}{\tau_\beta} l_{mn} x_{\alpha n} \tau_\alpha\right] \\ &= \sum_{\alpha,\beta=1}^{r} \left[\sum_{m,n} x_{\alpha m} l_{mn} x_{\beta n}\right]^2 \\ &\leqq N, \text{ independent of } r, \end{aligned}$$

since $\sum_{m,n} l_{mn}^2$ converges and $x_{\alpha m}$ is bounded.

Therefore the system of equations

(14) $$m_{\alpha i} = \xi_\alpha \sum_\beta c_{i\beta} m_{\alpha\beta}$$

has solutions $m_{\alpha i}$ of finite norm for real ξ_α, such that

(15) $$c_{i\beta} = \sum_\alpha \frac{m_{\alpha i} m_{\alpha\beta}}{\xi_\alpha}.$$

Noting that $c_{i\beta} = c_{\beta i}$, and substituting in (14) the expression (13) for $c_{\beta i}$, we obtain

$$\frac{m_{\alpha i}}{\tau_i} = \xi_\alpha \sum_{\beta,m,n} \frac{x_{\beta m}}{\tau_\beta} l_{mn} x_{in} m_{\alpha\beta}.$$

For the present we *shall assume that* $\frac{m_{\alpha i}}{\tau_i}$ *is of finite norm.* Then

$$(16)\quad \sum_i \frac{m_{\alpha i}}{\tau_i} x_{i\nu} = \xi_\alpha \sum_{\beta,m} l_{m\nu} \frac{m_{\alpha\beta}}{\tau_\beta} x_{\beta n} - \xi_\alpha \sum_{i,\beta,m,n} x^*_{i\nu}\, x^*_{in}\, l_{mn} \frac{m_{\alpha\beta}}{\tau_\beta} x_{\beta m},$$

since $x^*_{i\nu}$ is orthogonal to the matrix $(t_{\alpha\beta})$ and hence orthogonal to x_{in}.

But

$$\sum_{m,n} x^*_{in}\, l_{mn}\, x_{\beta m} = \sum_{m,n,\nu} x^*_{in}\, l_{mn}\, \tau_\beta^2\, t_{m\nu}\, x_{\beta\nu} = \sum_{m,n,\nu} x^*_{in}\, t_{mn}\, l_{m\nu}\, \tau_\beta^2\, x_{\beta\nu} = 0.$$

Therefore (16) reduces to

$$(17)\quad \sum_i \frac{m_{\alpha i}}{\tau_i} x_{i\nu} = \xi_\alpha \sum_{\beta,m} l_{m\nu} \frac{m_{\alpha\beta}}{\tau_\beta} x_{\beta m}.$$

In (17) we have exhibited a solution $\{v_{\alpha m}\} = \{\sum_i m_{\alpha i}/\tau_i\, x_{im}\}$ of finite norm, under the above assumption that $m_{\alpha i}/\tau_i$ be of finite norm, of the equations (17). Since $\sum_{m,\nu} (l_{m\nu})^2$ converges, the adjoint system will also have solutions of finite norm, which we call $\{u_{\alpha\nu}\}$. Furthermore, it can be shown that

$$(18)\quad v_{\alpha m} = \sum_\nu t_{m\nu}\, u_{\alpha\nu}.$$

These solutions are expressed in terms of the solutions of the supplementary matrix $(c_{\alpha\beta})$ and the characteristic numbers of $(l_{\alpha\beta})$ are the same as those of $(c_{\alpha\beta})$.

THEOREM 2. *If there exists a symmetric matrix* $(t_{\alpha\beta})$ *satisfying conditions* (10) *for a matrix* $(l_{\alpha\beta})$ *such that* $\sum_{\alpha,\beta} (l_{\alpha\beta})^2$ *converges, the matrix* $(l_{\alpha\beta})$ *has real and only real characteristic numbers.*

With the question of the existence of the characteristic numbers settled, we turn to a consideration of an expansion for the matrix $(l_{\alpha\beta})$.

From (15) we have

$$(19)\quad \sum_{i,\beta} \frac{x_{i\nu}}{\tau_i} c_{i\beta} \frac{x_{\beta\delta}}{\tau_\beta} = \sum_\alpha \frac{\sum_i \frac{m_{\alpha i}}{\tau_i} x_{i\nu} \sum_\beta \frac{m_{\alpha\beta}}{\tau_\beta} x_{\beta\delta}}{\xi_\alpha}.$$

By substituting in (19) the expression (13) for $c_{i\beta}$, we obtain on the left hand side

$$\sum \frac{x_{i\nu}}{\tau_i} \frac{x_{im}}{\tau_i} l_{mn}\, x_{\beta n}\, \tau_\beta \frac{x_{\beta\delta}}{\tau_\beta}.$$

And, as a result of property (10,d) the latter expression is equal to

$$\begin{aligned}\sum x_{i\nu}\, t_{m\nu}\, x_{ir}\, l_{mn}\, x_{\beta n}\, x_{\beta\delta} &= \sum t_{m\nu}\, l_{m\delta} - \sum x^*_{i\nu}\, x^*_{ir}\, t_{m\nu}\, l_{mn}\, x_{\beta n}\, x_{\beta\delta} - \sum t_{m\delta}\, l_{m\nu}\, x^*_{\beta n}\, x^*_{\beta\delta} \\ &= \sum t_{m\nu}\, l_{m\delta}.\end{aligned}$$

Therefore (19) reduces to

$$\sum_m t_{m\nu}\, l_{m\delta} = \sum_\alpha \frac{v_{\alpha\nu}\, v_{\alpha\delta}}{\xi_\alpha}. \tag{20}$$

THEOREM 3. *If there exists a symmetric matrix* $(t_{\alpha\beta})$ *satisfying conditions* (10) *for a matrix* $(l_{\alpha\beta})$ *such that* $\sum_{\alpha\beta}(l_{\alpha\beta})^2$ *converges, then*

$$\sum_m t_{m\nu}\, l_{m\delta} = \sum_\alpha \frac{v_{\alpha\nu}\, v_{\alpha\delta}}{\xi_\alpha}$$

where $\{v_{\alpha\nu}\}$ *represents a solution of the equations* (17) *and* ξ_α *a characteristic number of the equations* (17).

Applications to integral equations. We may now apply the method developed by Hilbert* for passing from a linear system of infinitely many equations in infinitely many unknowns to an integral equation.

Let

$$v(s) = \xi \int v(x)\, L(xs)\, dx, \tag{21}$$

and

$$u(s) = \xi \int L(sx)\, u(x)\, dx \tag{21'}$$

be any integral equation with continuous kernel, and its adjoint, respectively. Let $\int\int \varphi_i\, L\, \varphi_j = l_{ij}$, where $\{\varphi_i(s)\}$ represents a complete normed orthogonal system of functions. It is known that $\sum_{i,j}(l_{ij})^2$ converges.

Build up the functions

$$v_\alpha(s) = \xi_\alpha \sum_\nu v_{\alpha\nu} \int \varphi_\nu(x)\, L(xs)\, dx,$$

where $\{v_{\alpha\nu}\}$ is a solution of (17) for ξ_α, using as the matrix $(l_{m\nu})$ of (17) the $\{l_{m\nu}\}$ corresponding to the kernel L of (21), and

$$u_\alpha(s) = \xi_\alpha \sum_\nu u_{\alpha\nu} \int L(sx)\, \varphi_\nu(x)\, dx,$$

* Hilbert, "Grundzüge einer allgemeinen Theorie der linearen Integral-Gleichungen", pp. 186–188.

where $\{u_{\alpha\nu}\}$ is a solution of the adjoint of (17) for ξ_α. In the usual manner we prove that $v_\alpha(s)$ and $u_\alpha(s)$ are solutions of (21) and its adjoint respectively, for $\xi = \xi_\alpha$, and that these functions form a normed biorthogonal system.

THEOREM 4. *If $L(sx)$ is a continuous kernel such that there exists a symmetric matrix $(t_{\alpha\beta})$ satisfying conditions* (10) *for the matrix $(l_{\alpha\beta}) = \int\int \varphi_\alpha L \varphi_\beta$, where $\{\varphi_\alpha(s)\}$ represents a complete normed orthogonal system of functions, then the kernel $L(sx)$ has real and only real characteristic numbers.*

We consider next the expansion problems: (a) the determination of the class of functions $f(s)$ for which the series $\sum_\alpha v_\alpha(s) \int f u_\alpha$ converges; and (b) the convergence of this series to the function $f(s)$.

We shall say that Tg exists if $\sum_j t_{ij} g_j$ is of finite norm. And we shall write $f = Tg$ if $f_i = \sum_j t_{ij} g_j$ where $f_i = \int f \varphi_i$ and $g_i = \int g \varphi_i$.

In this discussion we shall need the

LEMMA. *The series $\sum_\alpha \left(\int g v_\alpha\right)^2$ converges for $g(s)$ such that Tg exists, or the series $\sum_\alpha \left(\int h u_\alpha\right)^2$ converges for $h = Tr$ where r is any continuous function such that Tr exists.* To prove this lemma, we note that because of (10, a)

$$\sum_{i,j=1}^{\infty} \left[g_i - \sum_{\alpha=1}^{n} u_{\alpha i} \int g v_\alpha\right] t_{ij} \left[g_j - \sum_{\alpha=1}^{n} u_{\alpha j} \int g v_\alpha\right] \geqq 0. \tag{22}$$

Under the given conditions upon g

$$\sum_{i,j=1}^{\infty} g_i t_{ij} g_j \leqq M.$$

After combining the remaining terms, (22) reduces to

$$M - \sum_{\alpha=1}^{n} \left(\int g v_\alpha\right)^2 \geqq 0,$$

which is the desired result.

From this lemma it follows that the series

$$\sum_\alpha \int f u_\alpha \int h u_\alpha \tag{23}$$

converges for $f = Tg$ and $h = Tn$, where g and n are any two continuous functions such that Tg and Tn exist. If we add the restriction $g(s) = \int L(sx)\, r(x)\, dx$, where r is any continuous function, we can prove that the series in (23) converges to $\int g\, Tn$. Under these conditions,

$$\sum_\alpha \int f u_\alpha \int h u_\alpha = \sum_\alpha \int g v_\alpha \int n v_\alpha = \sum \frac{1}{\xi_\alpha} \int r v_\alpha \int n v_\alpha.$$

But because of (20) this expression is equal to

$$\sum r_i t_{mj} l_{mi} n_j = \sum r_i l_{mi} h_m = \int\int h L r = \int g T n.$$

THEOREM 5. *If $g(s)$ and $n(s)$ are such that Tg and Tn exist and $g(s) = \int L(sx)\, r(x)\, dx$, where r is any continuous function*

$$\sum_\alpha \int g v_\alpha \int u_\alpha T n = \int g T n.$$

We can now discuss the expansion of an arbitrary function in terms of the solutions of (21) and its adjoint.

(a) The series

$$\sum_\alpha \frac{v_\alpha(s)}{\xi_\alpha^2} \int f u_\alpha \tag{24}$$

converges uniformly for $f = Tg$, where $g(s)$ is any continuous function such that Tg exists.

It is seen that the series in (24) converges uniformly when $(v_{\alpha m}/\xi_\alpha)$ is a bounded matrix and $\int f u_\alpha$ is of finite norm. If $f = Tg$, $\int f u_\alpha$ is of finite norm because of the above lemma. Also because of this lemma $\sum_{i,m} v_{\alpha i} l_{im} x_m$ is of finite norm, where x_m is any sequence of finite norm. Therefore the matrix $\sum_i v_{\alpha i} l_{im}$ is bounded.* However, since $\sum_i v_{\alpha i} l_{im} = v_{\alpha m}/\xi_\alpha$, the matrix $(v_{\alpha m}/\xi_\alpha)$ is bounded.

(b) The series in (24) converges to the function $\int\int f(x)\, L(xt)\, L(ts)\, dt\, dx$.

Let

$$j(s) = \sum_\alpha \frac{v_\alpha(s)}{\xi_\alpha^2} \int f u_\alpha - \int\int f(x)\, L(xt)\, L(ts)\, dt\, dx.$$

Then

$$\int j^2 = \sum_\alpha \frac{\int j v_\alpha \int f u_\alpha}{\xi_\alpha^2} - \int\int f L L j = \sum_\alpha \int\int T u_\alpha\, L L j \int f u_\alpha - \int\int f L L j.$$

And from (24)

$$\int j^2 = \int\int f L L j - \int\int f L L j = 0.$$

Therefore

$$j(s) \equiv 0.$$

* Hellinger and Toeplitz, "Theorie der unendlichen Matrizen", Math. Ann., 1910.

THEOREM 6. *If $g(s)$ is any continuous function such that Tg exists, and* $n(s) = \int\int (Tg)\,LL$,

$$n(s) = \sum_\alpha \frac{v_\alpha(s)}{\xi_\alpha^2} \int n u_\alpha,$$

and the series converges absolutely and uniformly.

Section 3. A Symmetrizing Matrix for the Kernel $K(sx) - p(s)\int A(t)\,K(tx)\,dt$.

We return now to the original problem, that is the discussion of equation (5). In order that we may apply the results of Section 2, we must show the existence of a matrix $(t_{\alpha\beta})$ satisfying conditions (10) and the convergence of $\sum_i (v_{\alpha i})^2$. It will be convenient to consider two cases separately, case 1: $p_i \neq 0$; and case 2: $p_i = 0$ for $i = m_j$, where $j = 1, 2, \cdots, n$.

Case 1. $a_i p_i \varkappa_i \geqq 0$, $p_i \neq 0$. Build up the symmetric matrix

$$(t_{\alpha\beta}) = \left(\frac{a_\alpha}{p_\alpha \varkappa_\alpha} e_{\alpha\beta}\right).$$

Properties (10, a), (10, b), and (10, d) are evidently satisfied. In order to show that (10, c) is satisfied, we write

$$l_{\nu\beta} = \frac{e_{\nu\beta}}{\varkappa_\nu} - \frac{a_\beta p_\nu}{\varkappa_\beta}.$$

Then

$$\sum_\nu t_{\alpha\nu} l_{\nu\beta} = \frac{a_\alpha}{p_\alpha \varkappa_\alpha^2} e_{\alpha\beta} - \frac{a_\alpha a_\beta}{\varkappa_\alpha \varkappa_\beta}$$
$$= \sum_\nu t_{\beta\nu} l_{\nu\alpha}.$$

From its form it is also evident that $(t_{\alpha\beta})$ has only a point spectrum.

Case 2. $a_i p_i \varkappa_i \geqq 0$, $p_i = 0$ for $i = m_j$, $j = 1, 2, \cdots, n$. Assume

$$E_{m_j} = \sum_i{}' \frac{a_i p_i \varkappa_i}{\varkappa_i - \varkappa_{m_j}} \neq 0 \text{ for } j = 1, 2, \cdots, n. \tag{25}$$

Build up the symmetric matrix

$$(t_{\alpha\beta}) = \begin{cases} \left(\dfrac{a_\alpha}{p_\alpha \varkappa_\alpha} e_{\alpha\beta}\right) & \alpha \neq m_i,\ \beta \neq m_j \\ \left(\dfrac{a_\alpha a_\beta}{(\varkappa_\alpha - \varkappa_\beta) E_\beta}\right) & \alpha \neq m_i,\ \beta = m_j \\ \left(\dfrac{a_\alpha a_\beta}{E_\alpha E_\beta} \sum_r{}' \dfrac{a_r p_r \varkappa_r}{(\varkappa_r - \varkappa_\alpha)(\varkappa_r - \varkappa_\beta)}\right) & \alpha = m_i,\ \beta = m_j. \end{cases}$$

It is not so easily seen that properties (10) are satisfied in this case, and that $(t_{\alpha\beta})$ has only a point spectrum. However, we obtain by direct substitution and the proper arrangement of terms

$$\sum_{\alpha,\beta} x_\alpha\, t_{\alpha\beta}\, x_\beta = \sum_i \left[\sqrt{\frac{a_i}{p_i\, \varkappa_i}}\, x_i + \sum_{j=1}^{n} \frac{a_{m_j}\sqrt{a_i\, p_i\, \varkappa_i}}{E_{m_j}(\varkappa_i - \varkappa_{m_j})}\, x_{m_j}\right]^2,$$

which shows that

$$\sum_{\alpha,\beta} x_\alpha\, t_{\alpha\beta}\, x_\beta \geqq 0,$$

the equality sign holding only for

$$x_i = -\, p_i\, \varkappa_i \sum_{j=1}^{n} \frac{a_{m_j}\, x_{m_j}}{E_{m_j}(\varkappa_i - \varkappa_{m_j})}, \qquad i = 1, 2, \cdots$$

that is, for n linearly independent linear forms.

To test for (10, b) we need consider only the convergence of the squares along an m_j row, since along the other rows we have but a finite number of terms. For such a row

$$\sum_\beta t_{\alpha\beta}^2 = \frac{a_\alpha^2}{E_\alpha^2} \sum_\beta{}' \frac{a_\beta^2}{(\varkappa_\beta - \varkappa_\alpha)^2} + \sum_{j=1}^{n} (t_{\alpha m_j})^2.$$

The infinite series on the right hand sides converges for every α, since a_β is of finite norm, and $\varkappa_\beta$ represents the characteristic numbers of a symmetric continuous kernel. Property (10, d) is obviously satisfied. For (10, c) we have three distinct cases.

For

$\alpha \neq \beta$, $\alpha \neq m_i$, $\beta \neq m_j$,
$$\sum_\nu t_{\alpha\nu}\, l_{\nu\beta} = \sum_\nu t_{\beta\nu}\, l_{\nu\alpha} \text{ as above;}$$

$\alpha = m_i$, $\beta \neq m_j$,
$$\sum_\nu t_{\alpha\nu}\, l_{\nu\beta} = \frac{a_\alpha\, a_\beta}{\varkappa_\alpha\, \varkappa_\beta}\left[\frac{\varkappa_\beta}{(\varkappa_\beta - \varkappa_\alpha)\, E_\alpha} - 1\right],$$
$$\sum_\nu t_{\beta\nu}\, l_{\nu\alpha} = \frac{a_\alpha\, a_\beta}{\varkappa_\alpha\, \varkappa_\beta}\left[\frac{\varkappa_\beta}{(\varkappa_\beta - \varkappa_\alpha)\, E_\alpha} - 1\right];$$

$\alpha = m_i$, $\beta = m_j$,
$$\sum_\nu t_{\alpha\nu}\, l_{\nu\beta} = \frac{a_\alpha\, a_\beta}{\varkappa_\alpha\, \varkappa_\beta}\left[-1 + \frac{1}{E_\alpha\, E_\beta} \sum_r \frac{a_r\, p_r\, \varkappa_r^2}{(\varkappa_r - \varkappa_\alpha)\,(\varkappa_r - \varkappa_\beta)}\right]$$
$$= \sum_\nu t_{\beta\nu}\, l_{\nu\alpha}.$$

The symmetrizing property of $(t_{\alpha\beta})$ is thus completely established.

In order to prove that $(t_{\alpha\beta})$ has only a point spectrum we shall not need the detailed form of the matrix. For this reason we write the system of equations with matrix $(t_{\alpha\beta})$ in the form

$$
\begin{aligned}
x_i &= \tau\left[t_{ii}\,x_i + \sum_{r=1}^{n} t_{im_r}\,x_{m_r}\right] && i = 1, 2, \cdots \\
x_{m_j} &= \tau\left[\sum_s t_{m_j s}\,x_s + \sum_{r=1}^{n} t_{m_j m_r}\,x_{m_r}\right] && j = 1, 2, \cdots, n,
\end{aligned} \tag{26}
$$

noting of course, that $t_{m_j i} = t_{j m_i}$.

This system may be reduced to the following form

$$
\begin{aligned}
x_{m_1} &= \tau\left[\sum_i \frac{t_{m_1 i}}{1-\tau\,t_{ii}}\,(\tau \sum_j t_{im_j}\,x_{m_j}) + \sum_{i=1}^{n} t_{m_1 m_i}\,x_{m_i}\right] \\
&\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
x_{m_n} &= \tau\left[\sum_i \frac{t_{m_n i}}{1-\tau\,t_{ii}}\,(\tau \sum_j t_{im_j}\,x_{m_j}) + \sum_{i=1}^{n} t_{m_n m_i}\,x_{m_i}\right],
\end{aligned}
$$

i. e. a system of n linear equations in n unknowns. To find the characteristic values we form the determinant of the coefficients. The system of equations (26) will have solutions only for values of τ which make the determinant vanish. The determinant is one of finite order, but the elements are expressed in terms of infinite series. Since the terms of these series are functions which are analytic except for the poles $\tau_i = 1/t_{ii}$ and since the series converge uniformly for any point which is not a pole, the series represent functions which are analytic except for these poles. Therefore the determinant itself is analytic except for poles. However the zeros of such a function are isolated in any region excluding these poles. For this reason the matrix $(t_{\alpha\beta})$ can have only a point spectrum.

We have then for case 2 a symmetric matrix $(t_{\alpha\beta})$, which satisfies properties (10) and which has only a point spectrum.

It will be recalled that in the formal discussion of the general theory in Section 2, we assumed the convergence of $\sum_\nu (\sum_i m_{\alpha i}/\tau_i\,x_{i\nu})^2$ where $\{m_{\alpha i}\}$ represents a characteristic solution of the equations (14), for $\xi = \xi_\alpha$, and τ_i^2 a characteristic number of the matrix $(t_{\alpha\beta})$. For the particular case under discussion, i. e. equation (5), this convergence can be proved as follows. From (17) we have

$$
\frac{1}{\xi_\alpha}\sum_i \frac{m_{\alpha i}}{\tau_i}\,x_{i\nu} - \frac{1}{\varkappa_\nu}\sum_\beta \frac{m_{\alpha\beta}}{\tau_\beta}\,x_{\beta\nu} = -\frac{a_\nu}{\varkappa_\nu}\,M_\alpha
$$

where M_α is independent of ν. Therefore, when $\xi_\alpha \neq \varkappa_\nu$

$$\sum_i \frac{m_{\alpha i}}{\tau_i} x_{i\nu} = -\frac{a_\nu}{\varkappa_\nu} \frac{M_\alpha}{\frac{1}{\xi_\alpha} - \frac{1}{\varkappa_\nu}}$$

and

$$\sum_\nu \left(\sum_i \frac{m_{\alpha i}}{\tau_i} x_{i\nu} \right)^2 = M_\alpha^2 \sum_\nu \frac{a_\nu^2}{\left[\frac{\varkappa_\nu}{\xi_\alpha} - 1 \right]^2}. \tag{27}$$

Since the characteristic numbers $\varkappa_\nu$ form a discret sequence, we may select the particular $\varkappa$, e. g. $\varkappa^*$ which is nearest to ξ_α. Then the right hand side of (27) will be

$$\leqq \frac{M_\alpha^2}{\left[\frac{\varkappa^*}{\xi_\alpha} - 1 \right]^2} \sum_\nu a_\nu^2 \leqq N_\alpha .$$

For $\varkappa_\nu = \xi_\alpha$, we have already indicated that the solution of (5) reduces to a solution of the equation with kernel K. But, since it is known that $\sum_i v_{\alpha i}^2$ converges for such solutions, it is not necessary to discuss the convergence for this case.

We have made no restrictions about the number of a_i which may be zero. It is not necessary to do so. If an infinite number of a_i are zero we shall have an infinite number of linear forms orthogonal to the matrix $(t_{\alpha\beta})$, which is allowed. On the other hand, if an infinite number of p_i are equal to zero and only a finite number of a_i are zero, we can build up a matrix similar to the one above, but which would operate on the other side of the matrix $(l_{\alpha\beta})$. It is to be noticed especially that the same constants E_i, defined in (25), serve, regardless of whether we build up the matrix $(t_{\alpha\beta})$ for the right or the left of the matrix $(l_{\alpha\beta})$. Thus we arrive at the following conclusions.

There exists a matrix $(t_{\alpha\beta})$ *with properties* (10) *if* $a_i p_i \varkappa_i \geqq 0$, *and if* $p_i \neq 0$ (*or* $a_i \neq 0$); *or if* $p_i = 0$ (*or* $a_i = 0$) $E_i \neq 0$ *only for* $i = m_j$ $j = 1, 2, \cdots, n$.

Under either of these sets of conditions, we have the following theorems.

THEOREM 7 (a). *Equation* (5) *has real and only real characteristic numbers.*

THEOREM 7 (b). *For* $f = Tg$ *where* $g(s)$ *is any continuous function, the series on the left of*

$$\sum_\alpha \frac{v_\alpha(s) \int f u_\alpha}{\xi_\alpha^2} = \int\int f(x)\, L(xt)\, L(ts)\, dt\, dx,$$

converges absolutely and uniformly to the function on the right, where $u_\alpha(s)$ *and* $v_\alpha(s)$ *are solutions of* (5) *and its adjoint respectively, for* $\xi = \xi_\alpha$

Section 4. Cases Excluded Above.

$\int Ap = 0$. Early in the discussion the condition $\int Ap \neq 0$ was imposed. It is interesting to see what happens when this condition is removed. If $\int Ap = 0$, for any solution u of (1) and (2) $\iint AKu = 0$. In this case we shall have

$$\mu = -\frac{\lambda \iint AK^{(2)} u}{\iint AKp}.$$

If we retain the condition $a_i p_i \varkappa_i \geqq 0$, $\iint AKp \neq 0$. Then in place of equation (4) we have

$$u(s) = \lambda \int \left[K(sx) - p(s) \frac{\iint A(t)\, K(tt_1)\, K(t_1 x)\, dt_1\, dt}{\iint AKp} \right] u(x)\, dx.$$

The matrix $(t_{\alpha\beta})$ which replaces the previous $(t_{\alpha\beta})$ for case 1 will be

$$(t_{\alpha\beta}) = \left(\frac{a_\alpha}{p_\alpha \varkappa_\alpha^2} e_{\alpha\beta} \right).$$

However in order that $(t_{\alpha\beta})$ be positive, $a_\alpha p_\alpha$ must be positive, and in this case we would not have $\int Ap = 0$. This means that we cannot have a situation analogous to the one above. It does show that for $\iint AKp \neq 0$, and for $a_i p_i \varkappa_i^2 \geqq 0$, we have fundamentally the same situation as above. For further extensions we have similar conclusions.

$a_i p_i \varkappa_i < 0$. It is interesting to note at this stage, that although $a_i p_i \varkappa_i \geqq 0$ presents itself as a sufficient condition in two different treatments, it is not a necessary condition that all the characteristic numbers be real. This can be shown by the following simple example.

$$a_1 = -1\left[-\frac{36}{23}\right] \qquad a_2 = \frac{1}{2}\left[-\frac{36}{23}\right] \qquad a_3 = \frac{1}{3}\left[-\frac{36}{23}\right]$$

$$p_1 = 1 \qquad p_2 = \frac{1}{2} \qquad p_3 = \frac{1}{3}$$

$$\varkappa_1 = 1 \qquad \varkappa_2 = 2 \qquad \varkappa_3 = 3.$$

The factor $-36/23$ is introduced so that $\sum_i a_i p_i = 1$, which is assumed in the development of the expression for which we have written $D(\lambda)$. In this case

$$D(\lambda) = -\frac{36}{23}\left[\frac{-1}{1-\lambda} + \frac{\frac{1}{2}}{2-\lambda} + \frac{\frac{1}{3}}{3-\lambda}\right].$$

$D(\lambda) = 0$ will be satisfied by the roots of the equation

$$\lambda^2 - 12\lambda + 23 = 0.$$

Since the discriminant of this equation is positive, $D(\lambda) = 0$ has two real roots. This illustration shows that $a_i p_i x_i \geqq 0$ is not a necessary condition for the reality of all of the roots of the Fredholm Determinant for the kernel $L(sx)$.

Other examples can be built up to show that in some cases we have imaginary roots, and in some cases no roots.

PARTITION POLYNOMIALS.*

By E. T. Bell.

A great variety of known and new arithmetical functions are included as special cases in the (finite) polynomials $P_n(z)$, where $z = (z_1, z_2, \cdots, z_s)$ is a matrix of s independent variables, defined in § 2. Although the polynomials, like many expressions in the theory of numbers, are ill adapted to numerical computation for even moderately large values of the integers involved, nevertheless it is interesting that such a diversity of functions should be united in one form which, from the point of view of algebra, is not only rational but also extremely simple. To illustrate the last remark it may be stated that all of the binary quadratic class number functions are very special cases of the polynomials, as also is the number of sets of integer solutions $x_j \gtreqless 0$ $(j = 1, \cdots, s)$ of

$$n = c_1 x_1^2 + c_2 x_2^2 + \cdots + c_s x_s^2,$$

where n is an integer, and c_j $(j = 1, \cdots, s)$ are integers > 0. This last number, I believe, has not hitherto been obtained exactly in any form.

1. **The polynomials ψ, Ψ.** Let $n > 0$ be an integer, and let C_j denote a set of distinct integers > 0. If C_j contains n we write $C_j | n$; if m divides n we write $m | n$. The set C_j may contain any finite number of elements (integers > 0), or an infinite number of elements, but no element may be repeated in the set. The last, it will appear as we proceed, is not a restriction on the generality of our functions.

The polynomials $\psi_n^{(j)}(z_j)$, for a given z_j, C_j, are defined by

$$\psi_n^{(j)}(z_j) = \sum d z_j^{\delta}, \tag{1}$$

where the sum extends to all pairs (d, δ) of integers $d, \delta > 0$ such that

$$n = d\delta, \qquad C_j | d; \tag{2}$$

the superscript (j) in (1) indicates the set C_j.

Consider now any sets C_j $(j = 1, \cdots, s)$. Let z_j $(j = 1, \cdots, s)$ be s independent variables and a_j $(j = 1, \cdots, s)$ arbitrary constants. The case of interest is that in which the a_j are integers not all zero, but this is not assumed until so stated. The sum

* Received June 6, 1927.

$$-\sum_{j=1}^{s} a_j \psi_n^{(j)}(z_j) \tag{3}$$

is uniquely known from (1), (2) when n and the matrices

$$z = (z_1, \cdots, z_s), \qquad C = (C_1, \cdots, C_s), \qquad a = (a_1, \cdots, a_s) \tag{4}$$

are given. Hence we may denote the sum (3) by either of

$$\Psi_n(z, C, a), \qquad \Psi_n(z), \tag{5}$$

the second form being sufficient when C, a are given by the context in a particular case.

To illustrate the several definitions we shall calculate three functions (5) in which $s = 1, 1, 2$ respectively, as these are of importance later. Let c be a constant integer >0, and let C_1, C_2 be the sets of all distinct odd, even multiples of c that are >0. Consider the sums (1) in which $z_1 = -1$ and C_1 is as just defined, also $z_2 = 1$ and C_2 as above;

$$\sum (-1)^{\delta} d, \qquad C_1 | d, \tag{6}$$

$$\sum d, \qquad C_2 | d. \tag{7}$$

Then the sum (6) is the sum of all divisors d of n that are odd multiples of c with even conjugates δ, diminished by the sum of all divisors d of n that are odd multiples of c with odd conjugates δ, while (7) is the sum of all divisors d of n that are even multiples of c. Write

$$n = 2^{\alpha} m, \qquad c = 2^{\beta} r, \qquad s = m/r, \qquad m, r \text{ odd},$$

and suppose first that $c | n$, so that s is an odd integer and $n = 2^{\alpha-\beta} sc$. Consider (6). Then δ is even only if $\alpha > \beta$; δ is odd only if $\alpha = \beta$. Hence, if $c | n$, the sum (6) has the value $c\sigma(s)$ or $-c\sigma(s)$, where $\sigma(s) =$ the sum of all the divisors of s, according as $\alpha > \beta$ or $\alpha = \beta$. Let

$$\sigma(x), \qquad \omega(x), \qquad \eta(x) \tag{8}$$

denote respectively the sum of all, of the odd, of the even divisors of x if x is an integer >0, and otherwise let each of (8) be zero. Then, from the above, we have the value of (6),

$$c(-1)^{n/c} \qquad \omega(n/c). \tag{9}$$

In a similar manner we see that the value of (7) is

$$c[\sigma(n/c) - \omega(n/c)]. \tag{10}$$

Since $\sigma(x) = \omega(x) + \eta(x)$, formulas such as (9), (10) can be written in several ways.

To illustrate (3), (5) take $a = (2,1)$ in (4). The resulting function is

$$c[\{1-2(-1)^{n/c}\}\,\omega(n/c) - \sigma(n/c)]. \tag{11}$$

As another example we consider s integers $c_j > 0\,(j = 1, \cdots, s)$, and construct from (11) the function

$$\sum_{j=1}^{s} c_j\,[\{1-2(-1)^{n/c_j}\}\,\omega(n/c_j) - \sigma(n/c_j)], \tag{12}$$

which is an important one for future theorems. The sum (12) is completely known when n and $c_j\,(j = 1, \cdots, s)$ are given.

The special case (13) of (12) in which $c_j = 1$ $(j = 1, \cdots, s)$ will also be required,

$$s[\{1-2(-1)^{n}\}\,\omega(n) - \sigma(n)], \tag{13}$$

in terms of which we express in § 5 the number of representations of n as a sum of s squares.

As a numerical example of (6), (9) take $n = 42$, $c = 3$. All the divisors of 42 that are odd multiples of 3 are 3, 21, which have the respective conjugates 14, 2. Hence, by (6), the value of (9) is $3 + 21 = 24$, while directly from (9) we get $3(-1)^{14}\omega(14)$, $= 3(1+7) = 24$. Again, for $n = 42$, $c = 6$, the odd multiples of 6 that divide 42 are 6, 42, whose conjugates are 7, 1; hence (6) now has the value $-6-42 = -48$, while (9) gives $6(-1)^{7}\omega(7) = -6(1+7) = -48$. For $n = 42$, $c = 7$, the even multiples of 7 that divide 42 are 14, 42, whose sum is 56; in this case (10) gives $7[\sigma(6) - \omega(6)] = 7[12-4] = 56$, checking (7), (10).

For $s = 4$ in (4) there is an extremely prolific function (from which we shall derive in § 4 the binary quadratic class number functions) constructed with

$$z = (-1, -1, 1, 1), \qquad C = (E, O, O, E), \qquad a = (a_1, a_2, a_3, a_4),$$

where E is the set of all even integers > 0, and O the set of all odd integers > 0. The corresponding sums (1) are therefore

$$\begin{aligned}
\psi_n^{(1)}(z_1) &= \sum d(-1)^{\delta}, && E|d;\\
\psi_n^{(2)}(z_2) &= \sum d(-1)^{\delta}, && O|d;\\
\psi_n^{(3)}(z_3) &= \sum d, && O|d;\\
\psi_n^{(4)}(z_4) &= \sum d, && E|d;
\end{aligned}$$

whence we find readily,

$$(14)\qquad \begin{aligned} &\psi_n^{(1)}(z_1) = -[1+(-1)^n]\,\omega(n), \quad \psi_n^{(2)}(z_2) = (-1)^n\,\omega(n),\\ &\psi_n^{(3)}(z_3) = \omega(n), \qquad\qquad\qquad \psi_n^{(4)}(z_4) = \eta(n), \end{aligned}$$

in which (by the given values of z, C), $z_1 = z_2 = -1$, $z_3 = z_4 = 1$, and the superscripts 1, 2, 3, 4 on the ψ refer to $C_1 = E$, $C_2 = O$, $C_3 = O$, $C_4 = E$ as the C_j in (1). Hence in this case for $a = (a_1, a_2, a_3, a_4)$, the function (5) is

$$(15)\qquad [a_1 - a_3 + a_4 + (-1)^n (a_1 - a_2)]\,\omega(n) - a_4\,\sigma(n),$$

where we have used $\eta(n) = \sigma(n) - \omega(n)$ to reduce the form given directly by (3). Thus, in the first of the notations (5), (15) is

$$(16)\qquad \Psi_n((-1, -1, 1, 1), \quad (E, O, O, E), \quad (a_1, a_2, a_3, a_4)).$$

2. **The polynomials** P. Let k_j $(j = 1, \cdots, n)$ be integers $\geqq 0$ such that, for $n > 0$ constant,

$$n = k_1 + 2k_2 + 3k_3 + \cdots + nk_n,$$

and let $\sum_k$ indicate a sum over all sets $k = (k_1, k_2, \cdots, k_n)$ satisfying this equation. Write for the moment, with $\Psi_j(z)$ $(j = 1, \cdots, n)$ as in (5),

$$\begin{aligned} \Psi_1^{k_1}(z)\,\Psi_2^{k_2}(z) \cdots \Psi_n^{k_n}(z) &\equiv N(k_1, k_2, \cdots, k_n),\\ (1^{k_1}\,2^{k_2} \cdots n^{k_n})\,(k_1!\,k_2! \cdots k_n!) &\equiv D(k_1, k_2, \cdots, k_n),\\ R(k_1, k_2, \cdots, k_n) &\equiv N(k_1, k_2, \cdots, k_n)/D(k_1, k_2, \cdots, k_n). \end{aligned}$$

Then the sum

$$(17)\qquad \sum_k R(k_1, k_2, \cdots, k_n)$$

is a polynomial in z_j $(j = 1, \cdots, s)$, which we shall call a *partition polynomial* of *rank* n, *argument* z, *set* C and *index* a, and denote by either of

$$(18)\qquad P_n(z, C, a), \quad P_n(z),$$

corresponding to (5). Evidently this polynomial is uniquely known when n, z, C, a are given. To $P_0(z)$, for all z, C, a, we assign the value 1.

Let x be a parameter, and indicate by $\prod_j$ a product with respect to all n_j such that $C_j | n_j$ $(j = 1, \cdots, s)$. Then the function of x,

$$\prod_1 (1 - z_1 x^{n_1})^{a_1} \prod_2 (1 - z_2 x^{n_2})^{a_2} \cdots \prod_s (1 - z_s x^{n_s})^{a_s}$$

is completely determined when z, C, a are assigned, and hence we may denote this function by either of

$$f(x,\ z,\ C,\ a),\quad f(x), \tag{19}$$

the latter being sufficient when z, C, a are clear from the context. By the equality of infinite series in what follows we mean equality only in the matric sense: the parameter x is treated as an indeterminate, and two power series in x are said to be equal when and only when the coefficients of like powers of x in the two series are equal for every power of x. Thus, if a given function of x be expanded by two different methods, the equality of the expansions is defined. Suppose that for given z, C, a the expansion of $f(x)$ has been obtained as a formal ($\equiv$ without attention to convergence) power series in x by Maclaurin's theorem. Then

$$f(x) = \sum_{n=0}^{\infty} P_n(z)\, x^n, \tag{20}$$

as may be seen from the indicated method of expansion on the right of

$$f(x) = \exp \log f(x). \tag{21}$$

In other words, $\exp \log f(x)$ is the generating function of (18), in the usual meaning of the term.

Addition of matrices $a = (a_1, \cdots, a_s)$, $b = (b_1, \cdots, b_s)$ is as usual, $a + b = (a_1 + b_1, \cdots, a_s + b_s)$. From (19), (20) we have

$$P_n(z, C, a+b) = \sum_{j=0}^{n} P_j(z, C, a)\, P_{n-j}(z, C, b), \tag{22}$$

an addition theorem for $P(z)$ with respect to its index, the set remaining constant and the argument the same in the three sequences of polynomials concerned.

Let

$$C_j = A_j + B_j, \qquad (j = 1, \cdots, s),$$

the indicated addition being logical. Then for $C = (C_1, \cdots, C_s)$ as before, and $A = (A_1, \cdots, A_s)$, $B = (B_1, \cdots, B_s)$, we have $C = A + B$ (matric addition) and the following addition theorem for $P_n(z)$ with respect to its set,

$$P_n(z, A+B, a) = \sum_{j=0}^{n} P_j(z, A, a)\, P_{n-j}(z, B, a). \tag{23}$$

If $C_j \mid A_j$ $(j = 1, \cdots, s)$ we write $C_j - A_j$ for the sets which remain when from C_j are deleted all elements of $A_j (j = 1, \cdots, s)$. Then, as in deriving (23) or directly from (23) since $C_j = C_j - A_j + A_j$, we have

$$\sum_{j=0}^{n} P_j(z, A, a)\, P_{n-j}(z, C-A, a) = P_n(z, C, a). \tag{24}$$

There is not a similar addition theorem with respect to the argument, but we have the following substitute. Let α_j, β_j $(j = 1, \cdots, s)$ be two sets of symbols, or ideal numbers, defined by the properties that each α_j, β_j is a definite one of 0, 1, and the value of $\alpha_j + \beta_j$ is 1 for $j = 1, \cdots, s$, so that α_j, β_j are not zero together, also that $\alpha_j n = n\alpha_j$ where n is an integer, and

$$\alpha_j n + \beta_j n = (\alpha_j + \beta_j) n \qquad (j = 1, \cdots, s).$$

If $(x_1, x_2, \cdots, x_s) = x$ is any matrix, denote by αx, βx respectively the matrices $(\alpha_1 x_1, \cdots, \alpha_s x_s)$, $(\beta_1 x_1, \cdots, \beta_s x_s)$. Then

$$P_n(z, C, a) = \sum_{j=0}^{n} P_j(\alpha z, \alpha C, \alpha a)\, P_{n-j}(\beta z, \beta C, \beta a). \tag{25}$$

3. **Elliptic partition polynomials.** As in (14) take $s = 4$ and $C_1 = C_4 = E$, $C_2 = C_3 = O$; as in (19) let $\prod_j$ refer to all n_j such that $C_j | n_j$, and write

$$Q_j(x) \equiv \prod_j (1 + x^{n_j}) \qquad (j = 1, 2);$$

$$Q_j(x) \equiv \prod_j (1 - x^{n_j}) \qquad (j = 3, 4);$$

$$T_1(x) \equiv 2x^{1/4} Q_4^8(x), \qquad T_2(x) \equiv 2x^{1/4} Q_1^2(x) Q_4(x),$$

$$T_3(x) \equiv Q_2^2(x) Q_4(x), \qquad T_4(x) \equiv Q_3^2(x) Q_4(x).$$

Let $\sum_n$, $\sum_m$ refer to all integers $n > 0$ and to all odd integers $m > 0$ respectively, and let $(a|b)$ be the Jacobi-Legendre symbol, so that $(-1|m) = (-1)^{(m-1)/2}$. Then from the classic elliptic product expansions we have

$$T_1(x^4) = 2\sum_m x^{m^2} m(-1|m), \qquad T_2(x^4) = 2\sum_m x^{m^2}.$$

$$T_3(x) = 1 + 2\sum_n x^{n^2}, \qquad T_4(x) = T_3(-x).$$

Let the a_{ji} denote s integers $\gtreqless 0$ whose sum $\neq 0$, the c_j integers > 0, the s_j $(j = 1, \cdots, 4)$ integers > 0 whose sum is s, and write

$$R_j(x) = \prod Q_j^{a_{ji}}(x^{c_{ji}}),$$

where the product extends to $i = 1, \cdots, s_j$. Then the $\psi_n(z)$ whose generating function is $R_1(x) R_2(x) R_3(x) R_4(x)$ will be called an *elliptic partition polynomial*. The following example leads to a great many interesting special cases.

Denote the function in (15), (16) by

$$\psi(n, a) \equiv \psi(n), \qquad a \equiv (a_1, a_2, a_3, a_4),$$

the form $\psi(n)$ sufficing when the value of a is stated in the context, and let $\Psi(n, a) \equiv \Psi(n)$ be the polynomial (17) or (18) when therein $\Psi_n(z) = \psi(n)$ $(n = 1, 2, \cdots)$. Then it follows from (19) that

$$Q_1^{a_1}(x)\, Q_2^{a_2}(x)\, Q_3^{a_3}(x)\, Q_4^{a_4}(x) = \sum_{n=0}^{\infty} \Psi(n, a) x^n. \tag{26}$$

Numerous theorems of Euler, Jacobi, Gauss and others on partitions give striking results for $\Psi(n, a)$ when a has certain special values, as also do many of the classic relations between the elliptic theta constants, of which we note only the following as illustrations. For a as in the left hand column below the values of $\psi(n)$ are as at the right. By (15) we have

$$\begin{aligned}
a &= (0, 0, 1, 1), && \psi(n) = -\sigma(n);\\
a &= (0, 0, 3, 3), && \psi(n) = -3\sigma(n);\\
a &= (0, 0, -1, 1), && \psi(n) = \omega(n) - \eta(n);\\
a &= (2, 0, 0, 1), && \psi(n) = 2[1 + (-1)^n]\,\omega(n) - \eta(n);\\
a &= (0, 2, 0, 1), && \psi(n) = -[2(-1)^n\,\omega(n) + \eta(n)];\\
a &= (0, 0, 2, 1), && \psi(n) = -[\omega(n) + \sigma(n)],
\end{aligned}$$

where ω, η, σ are as in (8). For these, $\Psi(n)$ has the following values. For the first, $\Psi(n) = 0$ or $(-1)^j$ according as n is not a pentagonal number or is the number $j(3j+1)/2$, $j \gtreqless 0$; for the second, $\Psi(n) = 0$ if n is not a triangular number, and is $(-1)^j\,(2j+1)$ if n is the number $j(j+1)/2$, $j \geqq 0$; for the third, $\Psi(n) = 0$ or 1 according as n is not or is a triangular number $\geqq 0$; for the fourth, $\Psi(n) = 0$ or 1 according as n is not the double of a triangular number or is $j(j+1)$, $j \geqq 0$; for the fifth, $\Psi(0) = 1$, $\Psi(n) = 0$ or 2 according as $n > 0$ is not or is the square of an integer > 0; for the sixth, $\Psi(0) = 1$, $\Psi(n) = 0$ or $2(-1)^n$ according as $n > 0$ is not or is the square of an integer > 0. To state one of these, say the third, more fully so that the nature of all may be evident, take in the notation of (17), $s = n$,

$$N(k_1, \cdots, k_n) = [\omega(1) - \eta(1)]^{k_1} \cdots [\omega(n) - \eta(n)]^{k_n},$$

$D(k_1, \cdots, k_n)$ as in (17), and $R(k_1, \cdots, k_n)$ in (17) constructed with those N, D. Then the polynomial (17), which is now a polynomial in $\omega(j)$, $\eta(j)$ $(j = 1, \cdots, n)$, has the value 0 or 1 according as n is not or is a triangular number $\geqq 0$.

The following is a curious theorem, whose elliptic equivalent will be easily recognized. Let $k > 0$ be an integer, and let the $\psi(n)$ for the following values of the matrix a,

$$(k, 0, 0, 0), \quad (0, k, 0, 0), \quad (0, 0, k, 0),$$

be $\psi_1(n)$, $\psi_2(n)$, $\psi_3(n)$ respectively, so that these are

$$k[1+(-1)^n]\,\omega(n), \quad -k(-1)^n\,\omega(n), \quad -k\,\omega(n),$$

and let the corresponding $\Psi(n)$ be $\Psi_j(n)$ $(j = 1, 2, 3)$. Then

$$2k\,\Psi_1(n-1) = \Psi_2(n) - \Psi_3(n) \qquad (n = 1, 2, 3, \cdots)$$

when and only when $k = 8$.

4. **Class numbers as elliptic partition polynomials.** For the customary definitions of the class number functions $E(n)$, $F(n)$ see H. J. S. Smith, *Report on the Theory of Numbers*, Art. 135; for the generating functions required in writing down from § 3 the following a, ψ, Ψ lists, see a paper by the writer in the *Bulletin of the American Mathematical Society*, vol. 30, p. 494, where

$$D(n) = [(-1)^n + \{1+(-1)^n\}\cos n\pi/2 + \{1-(-1)^n\}\sin n\pi/2]\,E(n)$$

was introduced. We shall not transcribe any of the formulas. For 9 values of a as follows we write down the $\psi(n)$ indicated,

(27)	(0, 0, 6, 3),	$-3[\omega(n)+\sigma(n)]$,
(28)	(6, 0, 0, 3),	$3[2(-1)^n\,\omega(n)-\eta(n)]$,
(29)	(0, 6, 0, 3),	$-3[2(-1)^n\,\omega(n)+\eta(n)]$,
(30)	(4, 0, 2, 3),	$2[1+2(-1)^n]\,\omega(n)-3\eta(n)$,
(31)	(0, 4, 2, 3),	$-2[1+2(-1)^n]\,\omega(n)-3\eta(n)$,
(32)	(2, 4, 0, 3),	$2[1-2(-1)^n]\,\omega(n)-3\eta(n)$,
(33)	(2, 0, 4, 3),	$[1+2(-1)^n]\,\omega(n)-3\sigma(n)$,
(34)	(0, 2, 4, 3),	$-[1+2(-1)^n]\,\omega(n)-3\sigma(n)$,
(35)	(4, 2, 0, 3),	$[7+2(-1)^n]\,\omega(n)-3\sigma(n)$.

From the generating functions of the several class number functions as given in the paper cited we write down the values of the elliptic partition polynomials $\Psi(n)$ constructed with the above $\psi(n)$, as follows:

$$(27), \quad 12(-1)^n\,E(n); \qquad (28), \quad \Psi(n) = 0 \text{ or } F(8j+3)$$

according as n is odd or $n = 2j$, $j \geqq 0$;

$$(29), \quad 12E(n); \qquad (30), \quad (-1)^n F(4n+2); \qquad (31), \quad 4D(n);$$
$$(32), \quad 2F(4n+1); \qquad (33), \quad (2|4n+1)\,F(4n+1);$$
$$(34), \quad 4(-1)^n D(n); \qquad (35), \quad F(4n+2).$$

For example, (27) states that if

$$N(k_1, \cdots, k_n) = (-3)^{k_1+\cdots+k_n}[\omega(1)+\sigma(1)]^{k_1} \cdots [\omega(n)+\sigma(n)]^{k_n},$$

and $D(k_1, \cdots, k_n)$ is as before, the $R(k_1, \cdots, k_n)$ in (17) with these N, D has the value $12(-1)^n E(n)$.

5. **Numbers of representations in certain quadratic forms.** From the two expressions for $T_s(x)$ in § 3, it is clear that if x be replaced by x^{c_j} $(j = 1, \cdots, s)$ and the resulting Q, T forms be compared, we have two different expressions for the number of sets of integer solutions of the equation in the introduction. Thus we obtain this number as the partition polynomial (17) constructed on the function (12) as $\Psi_n(z)$. Hence if $N(k_1, \cdots, k_n)$ is the product

$$\prod_{m=1}^{n}\left[\sum_{j=1}^{s}\{c_j(1-2(-1)^{m/c_j})\,\omega(m/c_j)-\sigma(m/c_j)\}\right]^{k_m},$$

and $D(k_1, \cdots, k_n)$ is as in (17), the sum (17) in which $R(k_1, \cdots, k_n)$ is constructed with these N, D is an explicit formula for the required number of sets of solutions.

In particular then the number of representations of n as a sum of s squares is the similarly constructed sum in which (13) replaces (12).

Similarly, if $s = m+r$, $m>0$, the number of representations of $n = 4j+m$, $j \geqq 0$, as a sum of s squares of which precisely m are odd and occupy the first s places in the representations, is the partition polynomial constructed as above on the function

$$2[m+(-1)^n(m-r)]\,\omega(n)-(m+r)\,\eta(n)$$

instead of (12) or (13).

If we use the known theorems concerning representations as sums of squares or in given forms $c_1x_1^2+\cdots+c_sx_s^2$, the above three general theorems give a large number of evaluations of particular partition polynomials as functions of the divisors of a single integer. Again, from the transformation theory of the theta functions may be derived an indefinite number of similar theorems in which the classes C_j in (2) are certain of the residue classes to any given modulus.

A NEW FORMULATION FOR GENERAL ALGEBRA.*

By John W. Young.

Introduction.

The concept of a general algebra has developed hitherto as a formal generalization from the 2-unit algebra ($x = a \cdot 1 + b \cdot i$) of ordinary complex numbers to an n-unit algebra ($x = \sum a_i u_i$), where the a_i are marks of a field F, in which two fundamental operations, addition and multiplication, have been defined (in addition to the "scalar multiplication" of elements of the algebra by marks of F). The transformations $x' = a + x$ and $x' = ax$ are then defined by these operations.

In Part I of the present paper it is proposed to lay a foundation for a conception of general algebra by taking sets of transformations on a general class S as fundamental and defining operations of the algebra in terms of them. The process is a direct generalization of one of the methods that may be used to for introducing analytic methods into projective geometry on the basis of geometric postulates, a problem which has its origin in von Staudt's algebra of throws.† The new point of view here developed is of interest because of the ease with which fundamental properties are derived and also because of the natural generalizations which it suggests, not only as to algebras with more than two fundamental operations and possible alternative laws defining the relations between these operations, but also as to possible generalizations of the usual laws. In this connection reference may be made to the generalization of the distributive law derived in § 7.

Part II is devoted to the more special problem of showing how a basis may be laid for the prevailing theory without the intervention of a field. For a more complete statement of this problem see § 11.

---

* Received June 3, 1927. The principal results of Part I of this paper were first presented to the American Mathematical Society on December 28, 1911 and April 26, 1913 (abstracts, Bulletin of the American Mathematical Society, vol. 18 (1912), pp. 227, 228 and vol. 19 (1913), p. 514); and on the basis of the present simpler set of assumptions on September 9, 1926. Part II was presented to the Society under the title *On the definition of the Scorza-Dickson algebras* on February 26, 1927.

† If P_0, P_1, P_∞ are any three distinct points of a straight line and A_a is the parabolic projectivity having P_∞ as double point and which transforms P_0 into P_a and M_a is the projectivity having P_0 and P_∞ as double points and which transforms P_1 into P_a, addition and multiplication are defined by the relations $P_a + P_b = A_a(P_b)$ and $P_a \cdot P_b = M_a(P_b)$. See Veblen and Young, *Projective Geometry*, vol. 1, p. 141 et seq.

Part I. The general formulation.

1. **The elementary algebraic system.** We consider a general class S, the elements of which are represented by $a, b, c, \cdots, x, y, \cdots$ and a set $\mathfrak{T}$ of transformations T on S; i. e., such that every T of $\mathfrak{T}$ associates with every element a of S some object denoted by $T(a)$. The system $\{S, \mathfrak{T}\}$ we call an *elementary algebraic system*, if it satisfies the following postulates:

1.1. *For every a of S and for every T of $\mathfrak{T}$, $T(a)$ is a uniquely determined element of S,*

1.2. *Corresponding to every element a of S there is determined a unique transformation T_a of $\mathfrak{T}$.*

1.3. *Corresponding to every transformation T of $\mathfrak{T}$ there is determined a unique element a of S, such that $T = T_a$.*

We now define a one-valued operation $\circ$ on ordered pairs of elements of S by the relation

1.4 $$a \circ b = T_a(b).$$

The operation $\circ$ is then, by virtue of postulates 1.1 and 1.2, defined for every ordered pair a, b of S; and, given such a pair, the result is a unique element of S. By postulate 1.3, if $a \circ x = b \circ x$ for every x in S, then $a = b$.

The transformation T_a is represented by the relation $x' = a \circ x$ as x ranges over S. The set $\mathfrak{T}$ is represented by this relation as a ranges over S.

2. **The adjoint transformations.** Associated with every a of S there is, in addition to T_a, a second transformation $\bar{T}_a$, called the *transformation adjoint to T_a*, defined by the relation

2.1 $$\bar{T}_a(x) = T_x(a) = x \circ a.$$

The set of all adjoint transformations $\bar{T}_a$ is called the adjoint set $\bar{\mathfrak{T}}$. It will be observed that $\bar{T}_a = T_a$ if, and only if, $a \circ x = x \circ a$ for every x of S. $\bar{T}_a$ is, of course, not in general in $\mathfrak{T}$. Further, $\bar{T}_a = T_a$ for every a of S if, and only if, $a \circ b = b \circ a$ for every pair a, b of S.

3. **Regular elements.** An element a of S such that the inverse transformation T_a^{-1} associates with every b of S a unique element $T_a^{-1}(b)$ of S is called a *regular element of S with respect to $\mathfrak{T}$*. The set of all such regular elements is a subclass of S which we denote by S_T. In particular cases every element of S may be regular ($S_T = S$), or S may contain no regular element ($S_T = 0$). Similar definitions are made with respect to the adjoint transformations $\bar{T}_a$. We thus obtain the notion of a *regular*

element with respect to $\bar{T}$, and that of the subclass $S_{\bar{T}}$ of all such regular elements. We may note that under these definitions $\mathfrak{T}$ *is simply transitive on* S_T *and* $\bar{\mathfrak{T}}$ *is simply transitive on* $S_{\bar{T}}$.*

If a is regular as to T, the equation $a \circ x = b$ has a unique solution x in S, and $T_x(a) = b$. If a is regular as to $\bar{T}$, the equation $x \circ a = b$ has a unique solution x in S, and $\bar{T}_x(a) = b$.

4. **The properties** I, R, In, $\bar{I}$, $\bar{R}$, $\bar{I}n$.

4.1. The set $\mathfrak{T}$ is said to have the *property* $I(\bar{I})$ if the identical transformation is in $T(\bar{T})$; i. e., if there exists in $T(\bar{T})$ a transformation $T_i(\bar{T}_{\bar{i}})$ such that, for every a of S, $T_i(a) = a$ $(\bar{T}_{\bar{i}}(a) = a)$.

If $\mathfrak{T}$ has the property I, $\bar{T}_a(i) = a$ for every a of S, thus defining the correspondence between S and $\bar{\mathfrak{T}}$. If $\mathfrak{T}$ has the property $\bar{I}$, $T_a(\bar{i}) = a$ for every a of S, thus defining the correspondence between S and $\mathfrak{T}$ (Postulates 1.2 and 1.3). *If* $\mathfrak{T}$ *has the properties* I *and* $\bar{I}$, *we have* $T_i(\bar{i}) = \bar{i} = i \circ \bar{i}$ *and* $\bar{T}_{\bar{i}}(i) = i = i \circ \bar{i}$ *and hence*

4.11 $$\bar{i} = i,$$

and

4.12 $$i \circ a = a \circ i = a \text{ for every } a \text{ in } S \quad (I, \bar{I}).$$

The element i *is then unique* and is called the *identical element of* S *as to* $\mathfrak{T}$.

4.2. The set $\mathfrak{T}$ is said to have property $R(\bar{R})$ if the resultant $T_a T_b$ $(\bar{T}_a \bar{T}_b)$ of any two transformations of $\mathfrak{T}(\bar{\mathfrak{T}})$ is in $\mathfrak{T}(\bar{\mathfrak{T}})$.

4.21. *If* T *has the property* R *and* $T_a T_b = T_c$, *we have*

$$a \circ (b \circ x) = c \circ x,$$

for every x *in* S, *where* c *is an element of* S *uniquely determined by the ordered pair* a, b. (*Generalized Associativity.*)

If T also has the property $\bar{I}$, we may place $x = \bar{i}$, and obtain $a \circ b = c$. Whence,

4.22. *If* $\mathfrak{T}$ *has the properties* R *and* $\bar{I}$, *and* a, b, x *are any three elements of* S, *we have* $a \circ (b \circ x) = (a \circ b) \circ x$. (*Ordinary Associativity.*)

By this last relation, we have at once $\bar{T}_{b \circ x}(a) = \bar{T}_x \bar{T}_b(a)$ for every a of S; hence the resultant of any two transformations of $\bar{\mathfrak{T}}$ is in $\bar{\mathfrak{T}}$. This gives

4.23. *The properties* R *and* $\bar{I}$ *imply the property* $\bar{R}$.

4.3. The set $\mathfrak{T}$ is said to have the property $In(\bar{I}n)$ if for every a of $S_T(S_{\bar{T}})$ the inverse $T_a^{-1}(\bar{T}_a^{-1})$ is in $\mathfrak{T}(\bar{\mathfrak{T}})$.

Obviously, if $S_T \neq 0$, the properties R, In imply I; and, if $S_{\bar{T}} \neq 0$, the properties $\bar{R}$, $\bar{I}n$ imply $\bar{I}$. Further,

* A set of transformations is said to be simply transitive on a class S if, given any two elements a, b of S, there exists a transformation of the set which transforms a into b, and for at least one pair a, b there is only one such transformation.

4.31. *Under assumptions R, In, $S_T \neq 0$, the set $[T_a]$ as a ranges over S_T forms a group.*

For, if a and b are in S_T, $T_a^{-1} = T_{a'}$ (say) and $T_b^{-1} = T_{b'}$ (say) are in S_T and a' and b' are in S_T, since the inverse of $T_{a'}$ is T_a and the inverse of $T_{b'}$ is T_b. Further, $T_{a'} T_{b'} = T_a^{-1} T_b^{-1} = (T_b T_a)^{-1}$ is in S_T. Hence, if a and b are in S_T and $T_b T_a = T_c$, c is in S_T. Under $\bar{I}$ we should then have $b \circ a = c$. Hence,

4.32. *Under assumptions R, In, $\bar{I}$, $S_T \neq 0$, if a and b are regular as to $\mathfrak{T}$, $b \circ a$ is regular.*

5. **The operation $\circ$ under properties** *R, In, $\bar{I}$, $S_T \neq 0$.* In what follows we assume properties *R, In,* $\bar{I}$ and $S_T \neq 0$. The operation $\circ$ is then associative (4.22). Further, if $T_{a'} = T_a^{-1}$ (a in S_T), from $T_a(x) = b$, follows $T_{a'} T_a(x) = T_{a'}(b) = x$. Hence,

5.1. *Under the above assumptions, if a is any regular element of S, the equation $a \circ x = b$ has a unique solution $x = a' \circ b$.*

We have seen that *R,* $\bar{I}$ imply $\bar{R}$ (4.23), that *R, In,* $S_T \neq 0$ imply *I* (4.3), and that *I,* $\bar{I}$ imply $\bar{i} = i$. (4.11). Hence we may conclude from the relation $T_a T_b = T_c$ that $a \circ b = c$. In particular, if $b = a'$, we have

5.2. *Under the above assumptions, if $T_{a'} = T_a^{-1}$, we have $a \circ a' = i = a' \circ a$.* Hence, if $T_{a'} = T_a^{-1}$, then $\bar{T}_{a'} = \bar{T}_a^{-1}$. Hence,

5.3. *The assumptions R, In, I, and $S_T \neq 0$, imply $\bar{I}n$.*

With a, a' having the same meanings as above, we have $x \circ a' \circ a = x$ for every x in S. Hence, $\bar{T}_a \bar{T}_{a'} = T_{\bar{i}}$. It follows that $\bar{T}_{a'}$ has an inverse $\bar{T}_a$ in $\bar{\mathfrak{T}}$, so that, if a is regular as to $\mathfrak{T}$, a is also regular as to $\bar{\mathfrak{T}}$, and conversely. We have then

5.4. *Under the assumptions R, In, $\bar{I}$ and $S_T \neq 0$, we have $S_T = S_{\bar{T}}$.*

By way of summary we may state.

5.5. *The assumptions R, In, $\bar{I}$ and $S_T \neq 0$ imply the properties, I, $\bar{R}$, $\bar{I}n$, and $S_T = S_{\bar{T}}$.*

6. **The system $\{S, \mathfrak{A}, \mathfrak{M}\}$. The properties** D, $\bar{D}$, D', $\bar{D}'$. We now consider a system $\{S, \mathfrak{A}, \mathfrak{M}\}$, where S is any class as before and $\mathfrak{A}$ and $\mathfrak{M}$ are two sets of transformations A and M, respectively, both subject to postulates 1.1, 1.2, and 1.3. The sets $\mathfrak{A}$ and $\mathfrak{M}$ each define an operation to which the preceding discussion would apply. Some relation connecting these two operations must however be made, if the operations defined by $\mathfrak{A}$ and $\mathfrak{M}$ are to be related. The theory of groups suggests as a very natural relation that one of the sets, say $\mathfrak{M}$, transform the other set, $\mathfrak{A}$, into itself; i. e., that for any M and A we have a relation of the form $MAM^{-1} = A'$, where A' is in $\mathfrak{A}$. Since, however, M^{-1} may not always exist in M, it is more general to define this relation in the form $MA = A'M$.

6.1. The set M is said to have the *property* D as to $\mathfrak{A}^*$ if for every a, b of S there exists a *unique* b_a such that $M_a A_b = A_{b_a} M_a$.

6.2. The set M has the *property* $\bar{D}$ as to $\mathfrak{A}$, if for every a, b of S there exists a *unique* $\bar{b}_a$ such that $\bar{M}_a A_b = A_{\bar{b}_a} \bar{M}_a$.

Under D we say that A_b is transformed into A_{b_a} by M_a and write $A_b \overset{M_a}{\to} A_{b_a}$ or more simply $A_b \to A_{b_a}$ when there is no danger of ambiguity. As b ranges over S, b_a will range over a subset S_a of S; A_{b_a} will range over a subset $\mathfrak{A}_a$ of $\mathfrak{A}$. Similarly, for $S_{\bar{a}}$ and $\mathfrak{A}_{\bar{a}}$. We have then

6.3 $$S_a \leqq S, \qquad S_{\bar{a}} \leqq S$$

6.4 $$\mathfrak{A}_a \leqq \mathfrak{A}, \qquad \mathfrak{A}_{\bar{a}} \leqq \mathfrak{A}.$$

Let A_c be another transformation of A: $M_a A_c = A_{c_a} M_a$. Then, by 6.1, we have

$$M_a A_b A_c = A_{b_a} M_a A_c = A_{b_a} A_{c_a} M_a.$$

Hence, if $A_b \to A_{b_a}$ and $A_c \to A_{c_a}$, then $A_b A_c \to A_{b_a} A_{c_a}$. If $A_c = A_b^{-1}$, then $A_{c_a} = A_{b_a}^{-1}$. These relations prove

6.5. *If the set* $\mathfrak{A}$ *contains a group, any such group is transformed into a group contained in* $\mathfrak{A}$.

7. **Addition and multiplication. The distributive Law.** Now let the operation defined by $\mathfrak{A}$ be called *addition* and denoted by $+$, and let the operation defined by $\mathfrak{M}$ be called *multiplication* and denoted by $\cdot$ (or by simple juxtaposition, $a \cdot b = ab$). If A has the property $\bar{I}$, let $\bar{i}$ be denoted by $\bar{0}$; and if A has the property In, let A_a^{-1} be denoted by A_{-a} (a regular as to $\mathfrak{A}$). The relation defining the property D (6.1) then gives, c being any element of S,

7.1 $$a(b+c) = b_a + (ac),$$

a *generalized form of distributivity* in which b_a depends only on b and a. Under assumption $\bar{I}$ we may set $c = \bar{0}$ in relation 7.1, which gives

$$ab = b_a + (a\bar{0}).$$

Under assumptions R and In, this gives $b_a = ab - (a\bar{0})$ ($a\bar{0}$ regular as to $\mathfrak{A}$). Whence we have

7.2 $$a(b+c) = (ab) - (a\bar{0}) + (ac) \quad \textit{Generalized Distributivity.}$$

* It would be more logical to indicate this relation by $D_{\mathfrak{A}}$ and, if a system with more than two operations is to be discussed, this would indeed be necessary. Since, however, we do not propose in this paper to discuss such an extension, no confusion will arise by the simpler notation.

If, for every a of S, $M_a(\overline{0}) = \overline{0}$, i. e., $a\overline{0} = \overline{0}$, this becomes

7.21 $a(b+c) = ab+ac$ *Ordinary Distributivity.*

7.3. If $\mathfrak{A}$ has the property $\overline{I}$ and $M_a(\overline{0}) = \overline{0}$ for every a in S, a set $\mathfrak{M}$ having the property D is said to have the property D' ($\overline{D}'$, if $\overline{M}_a(\overline{0}) = \overline{0}$ and M has the property $\overline{D}$).

7.4. *Multiplication is distributive on the left (in the ordinary sense) with respect to addition, if $\mathfrak{A}$ has the properties R, In, $\overline{I}$, and $\mathfrak{M}$ has the property D'.*

Similarly, we have

7.5. *Multiplication is distributive on the right (in the ordinary sense, $(b+c)a = ba+ca$) with respect to addition, if $\mathfrak{A}$ has the properties R, In, $\overline{I}$, and $\mathfrak{M}$ has the property $\overline{D}'$.*

8. **Commutative operations. The properties C, G and AG.** We have already observed (2.1) that the operation $\circ$ in a system $\{S, \mathfrak{T}\}$ is commutative, i. e., $a \circ b = b \circ a$ for every a, b in S, if and only if $T_a = \overline{T}_a$ for every a in S. If a set of transformations T has the latter property, it is said to have the property C. Further, properties R, In, $S_T = S$ imply that $\mathfrak{T}$ forms a group on S; $\mathfrak{T}$ is then said to have the property G. If it also has the property C, the group is abelian: $\mathfrak{T}$ is then said to have the property AG.

9. **The prevailing conception of a general algebra.** The prevailing conception of a general algebra (linear algebra, system of hypercomplex numbers) is that of a system $\{S, \mathfrak{A}, \mathfrak{M}\}$ with two operations, addition $(+)$ defined by $\mathfrak{A}$ and multiplication $(\cdot)$ defined by $\mathfrak{M}$, where $\mathfrak{A}$ has the property AG and $\mathfrak{M}$ has the properties D' and $\overline{D}'$. Such an algebra is associative, if multiplication is associative (in the ordinary sense); i. e., if $\mathfrak{M}$ has the property R' to the effect that $M_a M_b = M_{a \circ b}$ for every a, b in S (see 4.21). It is said to be commutative, if multiplication is commutative; i. e., if $\mathfrak{M}$ has the property C. It is called a division algebra, if $a \circ x = b$ and $x \circ a = b$ each have a unique solution x in S for every $a \neq 0$ (0 being the identity element with respect to addition) and every b in S. If S_0 denotes the class of elements of S distinct from 0, this implies that $\mathfrak{M}$ is a group on S_0. If $\mathfrak{M}$ has this property it is said to have the property G_0. To summarize, we have

9.1. *The prevailing conception of an algebra is that of a system $\{S, \mathfrak{A}, \mathfrak{M}\}$ in which $\mathfrak{A}$ has the property AG and $\mathfrak{M}$ has the properties D' and $\overline{D}'$. It is associative, if $\mathfrak{M}$ has the additional property R'; commutative, if $\mathfrak{M}$ has the property C. It is a division algebra, if $\mathfrak{M}$ has the property G_0.*

It should be observed that the above conditions are *necessary* to characterize what is meant in the current literature (for references see § 11, below)

by a general algebra; they are not necessarily sufficient. It will be shown in Part II of the present paper, however, that the above conditions are sufficient to derive the greater part of the current theory of finite algebras, if an assumption of finiteness is added. The above characterization differs from those current chiefly in the fact that no mention is made of a field.

10. **Illustrations.** To illustrate the preceding principles we may consider the problem of determining all the commutative algebras in which addition and multiplication are defined by abelian groups of collineations in the (complex, projective) plane. The only such groups that need be considered are two-parameter groups.* There are only three such groups:

I. *The group of all collineations leaving a triangle OPQ invariant;*

II. *The group of all collineations leaving two points P, Q and two lines PQ, PO invariant, the projectivities on PO being parabolic;*

III. *The group of all elations leaving every point of a line PQ invariant.*

If OPQ is taken as the triangle of reference and any element a of the algebra, i. e., any point a of the plane, is represented by the coördinates (a_1, a_2, a_3), the transformations T_a of these groups may be represented by the matrices:

$$\text{I.} \qquad T_a = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}, \qquad i = (1, 1, 1),$$

$$\text{II.} \qquad T_a = \begin{pmatrix} a_1 & 0 & 0 \\ a_2 & a_1 & 0 \\ 0 & 0 & a_3 \end{pmatrix}, \qquad i = (1, 0, 1),$$

$$\text{III.} \qquad T_a = \begin{pmatrix} a_3 & 0 & a_1 \\ 0 & a_3 & a_2 \\ 0 & 0 & a_3 \end{pmatrix}, \qquad i = (0, 0, 1).$$

The condition that the group $\mathfrak{M}$ chosen to define multiplication must transform into itself the group $\mathfrak{A}$ chosen to define addition, while every M must leave invariant the identity element of $\mathfrak{A}$ leads readily, by a consideration of the invariant configurations of the groups, to the fact that III is the only group available for defining addition and that either I or II may be used for defining multiplication. We thus obtain the two algebras:

$$1) \qquad \begin{aligned} a + b &= (a_3 b_1 + a_1 b_3,\ a_3 b_2 + a_2 b_3,\ a_3 b_3), & 0 &= (0, 0, 1) \\ a\,b &= (a_1 b_1,\ a_2 b_2,\ a_3 b_3), & 1 &= (1, 1, 1)\dagger \end{aligned}$$

* See § 3; a simply transitive group in the projective plane is a two-parameter group.

† We use 1 to denote the identity element with respect to multiplication.

2) $a+b$ as in 1),
$ab = (a_1 b_1, a_2 b_1 + a_1 b_2, a_3 b_3)$, $\quad 1 = (1, 0, 1)$.

The regular elements for addition are the points not on $PQ\,(x_3 \neq 0)$. The regular elements for multiplication are in case 1) the points not on a side of the triangle $OPQ\,(x_1 x_2 x_3 \neq 0)$; in case 2) the points not on the lines PO, $PQ\,(x_1 x_3 \neq 0)$. If we let S consist of all points not on PQ and use non-homogeneous coördinates ($x_3 = 1$), our definitions become:

1′) $a+b = (a_1+b_1, a_2+b_2)$, $\quad 0 = (0,0)$, $\quad S_A = S$,
$ab = (a_1 b_1, a_2 b_2)$, $\quad 1 = (1,1)$, $\quad S_M =$ points for which $x_1 x_2 \neq 0$.

2′) $a+b$ as in 1′),
$ab = (a_1 b_1, a_2 b_2 + a_1 b_2)$, $\quad 1 = (1,0)$, $\quad S_M =$ points for which $x_1 \neq 0$.

It may be noted that if in 1′) P and Q be taken as the circular points at infinity and S be confined to the real points of the plane not on PQ, algebra 1′) is equivalent to the algebra of ordinary complex numbers.

In a similar fashion, from all the three-parameter abelian groups of collineations in three-space the commutative algebras of this type could be derived; from a knowledge of all n-parameter abelian groups of collineations in S_n, the commutative algebras defined by collineations in S_n. Attention may, finally, be called to a set of algebras defined in S_n by equations analogous to 1′):

$a+b = (\cdots, a_i+b_i, \cdots)$, $i = 1, 2, \cdots n$, $0 = (0, 0, \cdots, 0)$, $S_A = S$,
$ab = (\cdots, a_i b_i, \cdots)$, $1 = (1, 1, \cdots 1)$, $S_M =$ points for which $x_i \neq 0$,

where the x_i denote any sort of non-homogeneous coördinates. In particular, for $n = 3$, the point x may be determined by the parameters x_i of the planes passing through x of three pencils of planes on lines u_i (skew or not). The special relations of the lines u_i to each other give various subcases. The transformations $x' = x+a$ and $x' = ax$ in these algebras are not in general collineations. The extension of these observations to S_n is obvious.* That these algebras are not as trivial as their defining equations seem to imply follows from the example of the algebra of ordinary complex numbers cited above.

Part II. Algebras $\{S, \mathfrak{A}^{AG}, \mathfrak{M}^{D', \overline{D}'}\}$.

11. **Introductory considerations.** If the properties AG, D', $\overline{D}'$ are stated in terms of the operations $+$ and $\cdot$ it is seen that an algebra of this type satisfies the following postulates:

* See J. W. Young and F. M. Morgan, *The Geometries associated with a certain system of Cremona groups*, Trans. Amer. Math. Soc., vol. 17 (1916), pp. 233–244; especially, for the above equations, p. 237.

11.1. *If a, b, c are any three elements of S,*

A 1. *$a+b$ is a uniquely determined element of S;*

A 2. $a+b = b+a$;

A 3. $(a+b)+c = a+(b+c)$;

A 4. *There exists in S an element 0 such that $a+0 = a$;*

A 5. *There exists in S an element $-a$ such that $a+(-a) = 0$;*

M 1. *ab is a uniquely determined element of S;*

M 2. $a(b+c) = ab+ac$;

M 3. $(b+c)a = ba+ca$.

11.2. In their definitions of an algebra S over a field F, Scorza* and Dickson† include among their postulates *A*1, *A*2, *A*3, *M*1, *M*2, *M*3, while *A*4 and *A*5 are theorems under the complete set of their postulates. The class of algebras we are here considering therefore certainly includes all the algebras defined by them. Their remaining postulates concern the definition of scalar multiplication, i. e., of products αx, where x is any element of S and α is any mark of the given field F, and the existence of a finite basis.‡

11.3. It is our purpose in the remainder of this paper to show that a large part, if not all, of the theory developed for an algebra over a field can be developed on the basis of postulates *A* and *M* stated above and a postulate of finiteness (to be stated later, § 17) without the intervention of a field. Fundamental in the conventional theory is the notion of a *linear set* of elements of S, i. e., of elements of S of the form $\alpha_1 x_1 + \cdots + \alpha_r x_r$, where the x_i are any given elements of S and the α_i range over the field F.§ Such a linear set is of order r, if the x_i are linearly independent. A linear system (x) of order 1 then consists of all the elements αx as α ranges over F. Clearly, every element x of S determines uniquely the linear set (x), which forms a group with respect to addition; and two such linear sets (x) and (y) either coincide or have no elements other

---

* G. Scorza, *Le algebre di ordine qualunque e le matrici di Riemann*, Rend. Circ. Mat. Palermo, vol. 45 (1921), pp. 1-204; the definition referred to is on p. 7. See also, G. Scorza, *Corpi Numerici e Algebre*, Messina, 1921, pp. 180, 181.

† L. E. Dickson, *Algebras and their Arithmetics*, Chicago, 1923, pp. 9, 10.

‡ Thus, in Dickson's definition (loc. cit.), while his sets of postulates I, IV are equivalent to our *A*1, *A*2, *A*3, *M*1, *M*2, *M*3, his set of postulates II includes $\alpha a = a\alpha$, $\alpha(\beta a) = (\alpha\beta)a$, $(\alpha a)(\beta b) = (\alpha\beta)(ab)$; III includes $(\alpha+\beta)a = \alpha a + \beta b$, $\alpha(a+b) = \alpha a + \beta b$; and V is to the effect that S shall contain a finite number of elements $v_1, \cdots, v_m$ such that every element of S can be expressed as a sum $\alpha_1 v_1 + \cdots + \alpha_m v_m$ of scalar products of $v_1, \cdots, v_m$ by marks $\alpha_1, \cdots, \alpha_m$ of F.

§ The importance for the theory of general algebras of this notion of a linear set (called by him a *complex*) was first established by Wedderburn in his fundamental paper, *On hypercomplex numbers*, Proc. Lond. Math. Soc., (2), vol. 5 (1908), pp. 77-118.

than 0 in common. Further, a linear set $(x_1, \cdots, x_r)$ of order r either contains completely any linear set (y) of order 1 or it has no element other than 0 in common with (y). These considerations lead to the definitions of *partition* and *L-partition* of a group given below (§ 13), on the basis of which our theory proceeds.

12. **Subgroups of S. The sum of two subgroups.** Any subclass B of S which forms a group with respect to addition will be called simply a subgroup of S. The following notation is familiar, but is given in the interest of clearness. If a subgroup C of S contains all the elements of a subgroup B, we shall write $C \geqq B$ or $B \leqq C$. If C contains B and also elements not in B, we write $C > B$ or $B < C$. If B and C are two subgroups, the elements common to B and C form a subgroup which is denoted by $B \wedge C$ or $C \wedge B$.

12.1. Let $B = [b]$ and $C = [c]$ be two subgroups, and form all the elements $b + c$. This class is denoted by $B + C$. $B + C$ is then a subgroup of S. If $B < C$, $C + B = C$, and conversely.

12.2. *Each element of the sum $B + C$ of two subgroups B and C of S is expressible in a single way as the sum of an element of B and of an element of C if, and only if, $B \wedge C = 0$.*

For, $b' + c' = b'' + c''$ if, and only if, $b' - b'' = c'' - c' = t$, where t is in $B \wedge C$.

12.3. If, for $r \geqq 2$, $P_1, \cdots, P_r$ are subgroups of S, we define their sum P by induction on r:

$$P = P_1 + P_2 + \cdots + P_r = (P_1 + \cdots + P_{r-1}) + P_r.$$

12.4. By 12.2 any element of P is expressible in a single way as a sum of r elements, one from each of the subgroups P_i, if, and only if, 0 is the only element common to $P_1 + \cdots + P_{j-1}$ and P_j for $j = 2, \cdots, r$. If these $r - 1$ relations hold, the subgroups P_i are said to be *independent*.

12.5. We may readily verify that if B, C, D are any three subgroups of S we have $(B + C) + D = B + (C + D)$. Hence, the sum $P_1 + \cdots + P_r$ has a unique meaning even without the definition by induction. It follows that if the $r - 1$ relations

$$(P_1 + \cdots + P_{j-1}) \wedge P_j = 0, \qquad j = 2, \cdots, r,$$

are satisfied, and if $i_1, \cdots, i_r$ are the numbers $1, \cdots, r$ in any other order, the $r - 1$ relations

$$(P_{i_1} + \cdots + P_{i_{j-1}}) \wedge P_{i_j} = 0$$

will also be satisfied.

13. **The partitions of a group. L-partitions.** A class $[H]$ of subgroups of any group G is called a *partition*, if every element a of G

is in one and only one subgroup H_a of $[H]$. Every group G has at least the partition consisting simply of G itself. If a partition contains at least two distinct subgroups it is called a *proper* partition. A partition $[H]$ is called an *L-partition*, if the group $\{H_1, \cdots, H_r\}$ generated by any finite number r of subgroups H is such that, if H_k is any other H, either $\{H_1, \cdots, H_r\}$ contains H_k or $\{H_1, \cdots, H_r\} \wedge H_k = 0$.*

14. **Linear sets in S.** Let $[H]$ be any L-partition of the group S with respect to addition, and let the subgroup H_a of S determined by any element a of S be called the *simple linear set* determined by a and be denoted by (a). A subgroup of S is called a linear set if it can be represented as the sum

$$(a)+(b)+\cdots+(p)$$

of a finite number of simple linear sets determined by $a, b, \cdots, p$. Such a linear set will be denoted by $(a, b, \cdots, p)$.

Linear sets as above defined satisfy the following conditions:

14.1. *If a and b are any two elements of S, $(a) = (b)$ if b is in (a).*

14.2. *If a and b are any two elements of S, $(a) \wedge (b) = 0$, if b is not in (a).*

14.3. *If b is any element of a linear set B, (b) is entirely contained in B;* i. e., $(b) \leqq B$.

14.4. *If b is not in a linear set B, we have $B \wedge (b) = 0$.*

15. **The fundamental theorem on linear sets.** We prove first:

15.1. *If P is a linear set $(a)+(b)$ with $(a) \wedge (b) = 0$ and Q is a linear set $(c)+(d)$ with $(c) \wedge (d) = 0$, then $P = Q$ if, and only if, c and d are in P*

If c and d are in P, we have $(c) \leqq P$ and $(d) \leqq P$ (14.3). Since P. is a group, every element of $(c)+(d)$ is in P. Hence, $Q \leqq P$. Now, let $d = b$ and $(a) \wedge (c) = 0$. If c is in P, let $c = a'+b'$. $Q' = (c)+(b)$ then contains $(a'+b')-b' = a'$ and, hence, contains (a) (14.1). Therefore, Q' contains every element of P; i. e., $Q' \geqq P$. By the same argument we show that $Q \geqq Q'$. Hence, $Q \geqq P$ and, in connection with the first result $Q = P$, if c and d are in P. If c (or d) is not in P, $(c) \wedge P = 0$ (or $(d) \wedge P = 0$) (14.4). Hence, in either case, Q contains elements not in P.

15.2. If the $r \geqq 2$ elements $p_1, \cdots, p_r$ of S satisfy the $r-1$ relations

$$(p_1, \cdots, p_{j-1}) \wedge (p_j) = 0, \qquad j = 2, \cdots, r,$$

they are said to be *linearly independent*. The linear set $(p_1, \cdots, p_r)$ is then said to be of order r. A simple linear set (p) defined by any one element $p(\neq 0)$ of S is of order 1. The zero set, consisting of 0 only, is of order 0. The theorem of (15.1) may now be stated as follows: *A*

* See a paper by the author, *On the partitions of a group and the resulting classification*, Bull. Amer. Math. Soc., vol. 33 (1927), pp. 453–461.

linear set of order 2 is uniquely determined by any two of its elements that are linearly independent. We may now prove the fundamental theorem:

15.3. *A linear set of order n is uniquely determined by any n of its elements that are linearly independent.*

The proof is by induction. We assume the theorem proved for $n = k$ and will then show that this assumption implies its validity for $n = k+1$. Since we have proved the theorem for the case $n = 2$, this will complete the proof.

Let $P_{k+1} = (p_1, \cdots, p_{k+1})$, where the p_i are linearly independent, be any linear set of order $k+1$. Let $q_1, \cdots, q_{k+1}$ be any $k+1$ linearly independent elements of P_{k+1}. We must prove $Q_{k+1} = (q_1, \cdots, q_{k+1}) = P_{k+1}$. To this end we observe first that not all the q_i can be contained in $P_k = (p_1, \cdots, p_k)$. For, if so, we should have $P_k = (q_1, \cdots, q_k)$ by the assumption that our theorem is true for $n = k$, and q_{k+1} would be contained in $(q_1, \cdots, q_k)$ contrary to the hypothesis that the q_i are linearly independent. There is then at least one q_i, say q_{k+1}, that is not in P_k. Let $Q' = (p_1, \cdots, p_k, q_{k+1})$. Since the basis of Q' is in P_{k+1} we have at once $Q' \leqq P_{k+1}$. Since q_{k+1} is in P_{k+1}, we have $q_{k+1} = p'_1 + \cdots + p'_{k+1}$, where p'_i is some element of (p_i), $i = 1, \cdots, k+1$. Q', therefore, contains the element $p'_{k+1} = q_{k+1} - (p'_1 + \cdots + p'_k)$, and hence Q' contains (p'_{k+1}). Now, $p'_{k+1} \neq 0$, for if it were, q_{k+1} would be in P_k. Hence, (p'_{k+1}) contains p_{k+1}, and Q' contains p_{k+1}. Hence, $Q' \geqq P_{k+1}$. Therefore, we have $Q' = P_{k+1}$.

We observe next that not all the elements $q_1, \cdots, q_k$ can be in $(p_1, \cdots, p_{k-1}, q_{k+1})$. Let q_k be not in $(p_1, \cdots, p_{k-1}, q_{k+1})$ and let $Q'' = (p_1, \cdots, p_{k-1}, q_k, q_{k+1})$. By reasoning entirely similar to that used above we prove $Q'' = Q' = P_{k+1}$. This argument may evidently be continued until every p_i in $(p_1, \cdots, p_{k+1})$ has been replaced by the corresponding q_i. In fact, at any stage of the process, if we have shown that $P_{k+1} = Q' = Q'' = \cdots = Q^{(j)}$, where

$$Q^{(j)} = (p_1, \cdots, p_{k-j+1}, q_{k-j+2}, \cdots, q_{k+1}).$$

Not all the $q_1, \cdots, q_{k-j+1}$ are in $(p_1, \cdots, p_{k-j}, q_{k-j+2}, \cdots q_{k+1})$; for, if they were, q_{k-j+1} would be in $(q_1, \cdots, q_{k-j}, q_{k-j+2}, \cdots, q_{k+1})$, contrary to the hypothesis that the q_i are independent. Let

$$Q^{(j+1)} = (p_1, \cdots, p_{k-j}, q_{k-j+1}, \cdots, q_{k+1}).$$

As before, we have $Q^{(j)} \geqq Q^{(j+1)}$. Also since q_{k-j+1} is in $P_{k+1} = Q^{(j)}$, we have

$$q_{k-j+1} = p_1^{(j+1)} + \cdots + p_{k-j+1}^{(j+1)} + q_{k-j+2}^{(j+1)} + \cdots + q_{k+1}^{(j+1)},$$

from which it follows that $p_{k-j+1}^{(j+1)}$ is an element of $Q^{(j+1)}$, the latter, therefore, also containing p_{k-j+1}. Hence, we have $Q^{(j+1)} \geqq Q^{(j)}$; and, thence, $Q^{(j+1)} = Q^{(j)}$. The process then leads to the complete sequence of equalities

$$P_{k+1} = Q' = \cdots = Q^{(j)} = \cdots = Q^{(k+1)} = Q_{k+1}.$$

16. **The product of linear sets.** The product PQ of two linear sets P and Q of an algebra is defined in the usual way as the linear set of minimum order which contains all elements obtained by multiplying each element of P by each element of Q. Let u and v be any two elements of S and consider the simple linear sets (u) and (v). If for every choice of u and v we have $(u)(v) = (uv)$, the partition of $\mathfrak{A}$ defining (u) and (v) is said to be *invariant.*

We may now prove in the usual way, that if the L-partition defining the simple linear sets is invariant, *the order of the product of two linear sets is equal to or less than the product of their orders.*

16.1. An abelian group $\mathfrak{A}$ admits a proper partition if, and only if, all of its elements are of the same prime order p or if it contains no elements of finite order.* The socalled primitive partition of such a group consists in the first case of the cyclical subgroups of order p of $\mathfrak{A}$, in the second case of the subgroups H_a, where H_a is the smallest subgroup of $\mathfrak{A}$ containing all cyclical subgroups containing the element a. The necessary and sufficient condition that the primitive partition of an abelian group $\mathfrak{A}$ containing no elements of finite order be an L-partition is that for every element A_s of $\mathfrak{A}$ and any integer k there exists in $\mathfrak{A}$ an element A_t such that $A_t^k = A_s$.†

16.2. On the basis of these results we may prove that a proper primitive L-partition of the additive group $\mathfrak{A}$ of an algebra is always invariant.

If any element A_t of $\mathfrak{A}$ is represented by the corresponding element t of S, the successive powers $A_t^2, A_t^3, \cdots, A_t^n$, of A_t are $t+t$, $t+t+t$, $\cdots$, $t+t+\cdots$ to n terms, and these are conveniently represented by $2t$, $3t$, $\cdots$, nt, respectively.

If all the elements of $\mathfrak{A}$ are of the same prime order p and a, b are any two elements of $\mathfrak{A}$ and (a) and (b) the two simple linear sets defined for a and b by a proper primitive partition of $\mathfrak{A}$, the elements of (a) are ma, and the elements of b are nb, where m, n are marks of a modular field of order p (i. e., $pa = 0$). Any element of $(a)(b)$ is then $(ma)(nb)$.

* The proofs of this and the following theorems and the related definitions are found in the paper previously cited, J. W. Young, *On the partitions of a group etc.,* loc. cit., pp. 459

† The primitive partition of an abelian group all of whose elements are of the same prime order is always an L-partition.

But the distributive law yields readily that $(ma)(nb) = (mn)(ab)$, and when m and n range independently over the modular field mn ranges over this field. Hence, we have $(a)(b) = (ab)$.

If $\mathfrak{A}$ contains no elements of finite order and admits a proper L-partition, corresponding to every s of $\mathfrak{A}$ and every integer n, there exists in $\mathfrak{A}$ an element t such that $nt = s$, or symbolically, such that $t = (1/n)s$. The group H_s of a proper primitive L-partition of $\mathfrak{A}$ then consists of all elements $(m/n)s$, where m, n are any integers $(n \neq 0)$. The distributive law again shows that, if u and v are any elements of A, $(m/n)u(q/r)v = (mq/nr)(uv)$. This proves that in this case also $(u)(v) = (uv)$.

17. **The postulate of finiteness.** We may now state our postulate of finiteness as follows:

F. *With respect to some invariant L-partition of* $\mathfrak{A}$, *the algebra has a finite basis.*

This means that the algebra S shall contain a finite number of elements $u_1, \cdots, u_n$ such that every element of S is in the linear set $(u_1, \cdots, u_n)$ where the simple linear sets are the groups of the invariant L-partition.

18. **Conclusion.** On the basis of postulates A, M, and F (see §§ 11 and 17) the prevailing theory of algebras may now be developed by the usual methods. An invariant L-partition serves to define the linear sets, in terms of which the theory is built up, and the fundamental theorems on linear sets have been derived in the preceding pages. Of course, a postulate of associativity for multiplication must be added where the prevailing theory requires it.

Thus, it appears that the first twenty-four theorems of Wedderburn's fundamental paper follow precisely as proved by him. A similar statement may be made concerning Scorza's treatment. The first 43 sections of Dickson's book also may be derived on the above basis.* The first difficulty seems to arise in connection with the rank equation; but it is not at all clear that the present formulation may not be adapted to this topic as well. In any case enough has been said to show that the explicit postulation of a field in connection with the development of the prevailing theory of algebras is to a very considerable extent unnecessary. Precisely to what extent, if at all, the present formulation is more general than the current one in terms of a field is a problem of no little interest.

DARTMOUTH COLLEGE,
HANOVER, N. H.

* It is necessary in a few instances, however, to use alternative proofs; e. g., for the theorem that every algebra which is not nilpotent contains an idempotent element (loc. cit., p. 44), for the proof of which Wedderburn's proof may be substituted.

A DETERMINATION OF THE GROUPS OF ORDER p^5.

By H. A. Bender.*

1. **Introduction.** A determination of the groups of order p^5 is given by G. Bagnera in Annali di Matematica, ser. 3, vol 1 (1998), pp. 137–228. In this discussion 69 pages have been devoted to the determination of the groups of order p^5, and the remaining 23 pages to the determination of the groups of order p^3 and p^4. One group for $p = 3$ has been omitted, and for $p = 3$ two groups which are simply isomorphic are listed as distinct groups. Prof. G. A. Miller, in the Bulletin of the American Mathematical Society, vol. 5 (1899), called attention to the incorrect† number of groups of order 32.

In the above determination of the groups of order p^5 they were studied in the most general manner without observing any of their special properties. By making use of some of these properties the work is greatly reduced and simplified, and at the same time many properties of the individual group are brought out.

Since every group of order p^m, $m > \beta(\beta-1)/2$, contains an invariant abelian subgroup of order p^β, it follows that every group of order p^5 contains an invariant abelian subgroup of order p^3. We shall study the groups of order p^5 under two classes. In the first class we shall include all the groups which contain an abelian subgroup of order p^4. In the second class we shall include all the groups which contain an abelian subgroup of order p^3, but which do not contain an abelian subgroup of order p^4. In each class we shall suppose p to be an odd prime number.

2. **Groups of order p^5 which contain an abelian subgroup of order p^4.** We shall represent the abelian subgroup of order p^4 by G_4, and suppose S_5 to be an operator which will with G_4 generate a group G_5 of order p^5. Since the pth power of every operator in the major co-set $G_4 S_5$ must be in G_4,‡ we see from the well known formula

$$S_m^{-p} S_a S_m^p = S_{a-p} S_{a-p+1}^p \cdots S_{a-1}^p S_a, \tag{1}$$

* Received June 29, 1927.

† This was corrected by G. Bagnera, Annali di Mat., (3), vol. 2, p. 263.

‡ Amer. Journ. Math., vol. 45 (1923), p. 231.

where $S_m^{-1} S_\beta S_m^p = S_{\beta-1} S_\beta$ for $(\beta = \alpha, \alpha-1, \cdots, \alpha-p+1)$, that it is necessary for

$$S_{\alpha-p} S_{\alpha-p+1}^p \cdots S_{\alpha-1}^p = 1. \tag{2}$$

Since $\alpha \leqq 4$, it is at once evident from (2) that for $p > 3$ the commutator subgroups of the groups of order p^5 can not contain an operator of order p^2, and hence the subgroup which is generated by the pth powers of the independent generators is either the central, or is contained in the central.

The four most important steps in the determination of the groups of order p^5 which contain a given abelian group G_4 are: first, the number of ways the central can be selected; second, for a given central H the number of ways the commutator subgroup can be selected; third, for a given central H and a given commutator subgroup K the possible ways G_4 can be transformed by the operators of the major co-set $G_4 S_5$; fourth, the total number of groups which transform G_4 in a given manner.

Since two groups are distinct if their centrals are not isomorphic, it follows that two groups are distinct if their centrals do not contain the same number of independent generators of like order. Two groups with a given central are distinct if their commutator subgroups do not include the same number of operators which are generated by the invariant independent generators, for in this case the commutator subgroups could not be made simply isomorphic. The third step consists of determining the number of distinct ways G_4 can be transformed when the groups contain a given central and a given commutator subgroup. It should be observed that two groups are distinct if the commutators formed by S_5 and the independent generators of like orders are not simply isomorphic.

Thus far two groups are considered to be distinct if their centrals, or their commutator subgroups, or the subgroups which are generated by the commutators formed by S_5 and the independent generators of like orders are not simply isomorphic. We have yet to determine those groups which have a given central and a given commutator subgroup, and in which the subgroups which are generated by the commutators formed by S_5 and the independent generators of G_4 which are of the same order, are isomorphic. Since this condition is possible only in case the commutator subgroup contains non-invariant operators, and operators which are generated by non-invariant independent generators, we shall treat each such case separately.

The fourth step consists in determining the number of distinct groups which transform G_4 in the same manner, that is, to determine the possible operators which may be the pth power of S_5. The pth power of any operator $S_5 S_\alpha$ of the major co-set $G_4 S_5$ is

$$(3) \qquad (S_5\, S_\alpha)^p = S_5^p \cdot S_{\alpha-p+1}\, S_{\alpha-p+2}^p \cdots S_\alpha^p.$$

Since for $p>3$ the commutators are all of order less than p^2, it follows that

$$(4) \qquad (S_5\, S_\alpha)^p = S_5^p\, S_\alpha^p.$$

Hence for $p>3$ any group of order p^5 which is generated by G_4 and S_5 is conformal with the abelian group which is generated by G_4 and an operator of the same order as S_5. Furthermore, if in any group of order p^5 $(p>3)$ the major co-set $G_4\, S_5$ contains an operator whose pth power is in the subgroup generated by the pth powers of the independent generators of G_4, then *this major co-set must contain an operator of order p.* Hence *we may suppose S_5 to be an operator of order p, or to be such that its pth power is an invariant independent generator of G_4.*

In what follows we shall make use of the following theorem: *If a group G of order p^m, p being an odd prime number, contains an abelian subgroup G_{m-1} of order p^{m-1} generated by λ $(\lambda<p-1)$ independent generators S_1, S_2, S_3, $\cdots$, S_λ of order p^{α_1} $(\alpha_1>1)$ such that the group generated by any number of these generators has only the identity in common with the group generated by the remaining generators, and a subgroup generated by less than λ generators is not invariant under G, nor is any cyclic subgroup of order p^{α_1} invariant under G, then all the operators in the major co-set $G_{m-2}\, S_1$ are conjugate under the group of isomorphisms of G.*

To prove this we may suppose $S_m^{-1} S_i S_m = s_{i+1} S_i$ $(i=1, 2, \cdots, \lambda-1)$ and $S_m^{-1} S_\lambda S_m = s_1^{x_1} s_2^{x_2} \cdots s_\lambda^{x_\lambda} S_\lambda$ where $S_i^{p^{\alpha_1-1}} = s_i$, S_m being an operator not in G_{m-1}. We shall prove that the following isomorphism is possible. Let S_m correspond to itself, S_1 to $S_1^{a_1} S_2^{a_2} \cdots S_\lambda^{a_\lambda}$, and S_j to $S_1^{a_{1j}} S_2^{a_{2j}} \cdots S_\lambda^{a_{\lambda j}}$ $(j=2, 3, \cdots, \lambda)$ where it is necessary for $a_{ij} = a_{i-j+1} + A_\lambda x_{i-j+2} + A_{\lambda-1} x_{i-j+3} + \cdots + A_{\lambda-j+2} x_i$ $(A_i = a_i + A_\lambda x_{i+1} + A_{\lambda-1} x_{i+2} + \cdots + A_{i+1} x_\lambda)$ $(a_n = 0, x_n = 0,\ n \leqq 0)$ $(i=1, 2, \cdots, \lambda)$.

Such an isomorphism is possible if

$$a_i x_1 + \sum_{j=2}^{\lambda} x_j a_{ij} = \sum_{k=1}^{i} A_{i-k+1} x_k.$$

It is easy to verify that this is a true relation by expanding and factoring x_k out of the left hand member.

In case there exists an n such that

$$n^\lambda x_1 + n^{\lambda-1} x_2 + \cdots + n x_\lambda - 1 \equiv 0 \bmod p,$$

then G contains a cyclic invariant subgroup generated by $S_1^{nx_1} S^{n^2 x_1 + n x_2} \cdots S_\lambda$.

Hence S_1 is conjugate, under the group of isomorphisms of G, to all the operators in the major co-set $G_{m-2} S_1$ except those operators which might be transformed into some power of themselves. Thus if two groups in which $S_m^{-1} S_i S_m = s_{i+1} S_i$ $(i = 1, 2, \cdots, \lambda - 1)$, and in the one group $S_m^{-1} S_\lambda S_m = s_1^{x_1} s_2^{x_2} \cdots s_\lambda^{x_\lambda} S_\lambda$ and in the other group $S_m^{-1} S_\lambda S_m = s_1^{y_1} s_2^{y_2} \cdots s_\lambda^{y_\lambda} S_\lambda$ are isomorphic they can be made isomorphic by making S_1 isomorphic to itself. Furthermore, since these groups contain but one abelian subgroup of order p^{m-1} if the commutator subgroup is of order greater than p, then, we may in addition make S_m isomorphic with some power of itself.

Hence all groups of order p^5 which are isomorphic and contain an abelian subgroup of order p^4 $(p > 3)$, can be made isomorphic by making an arbitrary operator of the abelian subgroup isomorphic to itself if the commutator subgroup is of order greater than p.

For each group we shall give the central H and the commutator subgroup K and the pth power of S_5. The manner in which G_4 may be transformed for a given H and K is at once evident, and we shall only give the transformations of G_4 in case there are more groups than the one which contain the same H and the same K for a given pth power of S_5. The manner of selecting H and K is an immediate consequence of the group of isomorphisms of G_4. We shall first give a determination of the groups of order p^4.

3. **Groups of order p^4.** The abelian groups of order p^4 are of types (1, 1, 1, 1), (2, 1, 1), (2, 2), (3, 1), and (4).

In the non-abelian groups of order p^4, first suppose G_3 is of type (1, 1, 1), and its generators are $S_1^p = S_2^p = S_3^p = 1$, then

1. $H = [S_1, S_2]$, $K = [S_1]$, $S_4^p = 1$; $S_4^p = S_1$ same as 4; $S_4^p = S_2$ same as 6;
2. $H = [S_1]$, $K = [S_1, S_2]$, $S_4^p = 1$; 3. $S_4^p = S_1$.

If G_3 is of type (2, 1), and its generators are $S_1^{p^2} = S_2^p = 1$, then

4. $H = [S_1^p, S_2]$, $K = [S_1^p]$, $S_4^p = 1$; 5. $S_4^p = S_2$;
6. $H = [S_1^p, S_2]$, $K = [S_2]$, $S_4^p = 1$; $S_4^p = S_2$ same as 5;
7. $H = [S_1]$, $K = [S_1^p]$, $S_4^p = 1$; $S_4^p = S_1$ same as 9;
8. $H = [S_1^p]$, $K = [S_1^p, S_2]$, $S_4^p = 1$, $S_4^{-1} S_1 S_4 = S_2 S_1$, $S_4^{-1} S_2 S_4 = S_1^{\alpha^n p} S_2$ $(n = 0, 1)$.*

* α a primitive pth root of unity modulo p.

If G_3 is of type (3), and its generator is $S_1^{p^3} = 1$, then

9. $H = [S_1^p]$, $K = [S_1^{p^2}]$, $S_4^p = 1$.

These groups are all conformal with abelian groups for $p > 3$. For $p = 3$ these groups are all conformal with abelian groups except groups 2, 3, and 8 ($n = 1$).

In 8 consider the isomorphism $\begin{pmatrix} S_4, & S_1, & S_2 \\ S_4^\mu, & S_1, & S_2^\mu \end{pmatrix}$ between the group in which $S_4^{-1} S_2 S_4 = S_1^{xp} S_2$, and the group in which $S_4^{-1} S_2 S_4 = S_1^{yp} S_2$. Hence the groups are isomorphic if the congruence

$$x \equiv y\mu^2 \mod p \tag{5}$$

is satisfied. From the fact $\mu^n \equiv 1 \mod p$ has d solutions, d being the greatest common divisor of n and $p-1$, it follows that there are only $(p-1)/d$ values of β for which there exists a μ which will satisfy the congruence

$$\beta\mu^n \equiv 1 \mod p. \tag{6}$$

Hence only those groups in which x and y are quadratic residues are isomorphic, and those in which x and y are quadratic non-residues modulo p are isomorphic.

4. **Groups of order p^5.** The abelian groups of order p^5 are of types (1, 1, 1, 1, 1), (2, 1, 1, 1), (2, 2, 1), (3, 1, 1), (3, 2), (4, 1), and (5).

In the non-abelian groups of order p^5, first suppose G_4 is of type (1, 1, 1, 1), and its generators are $S_1^p = S_2^p = S_3^p = S_4^p = 1$, then

1. $H = [S_1, S_2, S_3]$, $K = [S_1]$, $S_5^p = 1$; $S_5^p = S_1$ same as 9; $S_5^p = S_2$, or S_3 same as 11;
2. $H = [S_1, S_2]$, $K = [S_1, S_2]$, $S_5^p = 1$; 3. $S_5^p = S_1$, or S_2;
4. $H = [S_1, S_2]$, $K = [S_1, S_3]$, $S_5^p = 1$;
5. $S_5^p = S_1$, $(p > 3)$; 6. $S_5^p = S_2$;
7. $H = [S_1]$, $K = [S_1, S_2, S_3]$, $S_5^p = 1$ $(p > 3)$;
8. $S_5^p = S_1$, $(p > 3)$.

If G_4 is of type (2, 1, 1,), and its generators are $S_1^{p^2} = S_2^p = S_3^p = 1$, then

9. $H = [S_1^p, S_2, S_3]$, $K = [S_1^p]$, $S_5^p = 1$; 10. $S_5^p = S_2$, or S_3;
11. $H = [S_1^p, S_2, S_3]$, $K = [S_2]$, $S_5^p = 1$;

12. $S_5^p = S_3$, $S_5^p = S_2$ same as 10;

13. $H = [S_1, S_2]$, $K = [S_1^p]$, $S_5^p = 1$; $S_5^p = S_2$ same as 28; $S_5^p = S_1$ same as 34;

14. $H = [S_1, S_2]$, $K = [S_2]$, $S_5^p = 1$; $S_5^p = S_2$ same as 26; $S_5^p = S_1$ same as 35;

15. $H = [S_1^p, S_2]$, $K = [S_1^p, S_2]$, $S_5^p = 1$, $S_5^{-1} S_1 S_5 = S_2 S_1$, $S_5^{-1} S_3 S_5 = S_1^p S_3$; 16. $S_5^p = S_2$;

17. $H = [S_1^p, S_2]$, $K = [S_1^p, S_2]$, $S_5^p = 1$, $S_5^{-1} S_1 S_5 = S_1^p S_1$, $S_5^{-1} S_3 S_5 = S_2 S_3$; 18. $S_5^p = S_2$;

19. $H = [S_1^p, S_2]$, $K = [S_1^p, S_3]$, $S_5^p = 1$, $S_5^{-1} S_1 S_5 = S_3 S_1$, $S_5^{-1} S_3 S_5 = S_1^{\alpha' p} S_3$;

20. $S_5^p = S_2$ $(n = 0, 1)$; 20_1. $S_5^p = S_1^p$ $(n = 1)$ $(p = 3)$;

21. $H = [S_1^p, S_2]$, $K = [S_2, S_3]$, $S_5^p = 1$; 22. $S_5^p = S_2$;

23. $H = [S_1]$, $K = [S_1^p, S_2]$, $S_5^p = 1$; 24. $S_5^p = S_1$;

25. $H = [S_1^p]$, $K = [S_1^p, S_2, S_3]$, $S_5^p = 1$, $S_5^{-1} S_1 S_5 = S_3 S_1$, $S_5^{-1} S_3 S_5 = S_2 S_3$, $S_5^{-1} S_2 S_5 = S_1^{\alpha' p} S_2$ $(p > 3)$ $(n = 0)$; 25_1. if $p - 1 = 3k$ $(n = 1, 2)$.

If G_4 is of type (2, 2), and its generators are $S_1^{p^2} = S_2^{p^2} = 1$, then

26. $H = [S_1^p, S_2]$, $K = [S_1^p]$, $S_5^p = 1$; 27. $S_5^p = S_2$;

28. $H = [S_1^p, S_2]$, $K = [S_2^p]$, $S_5^p = 1$; 29. $S_5^p = S_2$;

30. $H = [S_1^p, S_2^p]$, $K = [S_1^p, S_2^p]$, $S_5^p = 1$, $S_5^{-1} S_1 S_5 = S_2^p S_1$, $S_5^{-1} S_2 S_5 = S_1^{\alpha' p} S_2^{\beta p} S_2 \left(\beta = 0, 1, \cdots, \frac{p-1}{2}\right)$ $(n = 0, 1)$;

31. $H = [S_1^p, S_2^p]$, $K = [S_1^p, S_2^p]$, $S_5^p = 1$, $S_5^{-1} S_1 S_5 = S_1^p S_1$, $S_5^{-1} S_2 S_5 = S_2^p S_2$;

32. $H = [S_1^3]$, $K = [S_1, S_2^3]$, $S_5^3 = 1$, $S_5^{-1} S_2 S_5 = S_1 S_2$, $S_5^{-1} S_1 S_5 = S_1^{6n} S_2^6 S_1$ $(n = 0, 1)$; 33. $S_5^3 = S_1^3$ $(n = 1)$.

If G_4 is of type (3, 1), and its generators are $S_1^{p^3} = S_2^p = 1$, then

34. $H = [S_1^p, S_2]$, $K = [S_1^{p^2}]$, $S_5^p = 1$; $S_5^p = S_2$ same as 29;

35. $H = [S_1^p, S_2]$, $K = [S_2]$, $S_5^p = 1$; $S_5^p = S_2$ same as 27;

36. $H=[S_1]$, $K=[S_1^{p^2}]$, $S_5^p=1$; $S_5^p=S_1$ same as 38;

37. $H=[S_1^p]$, $K=[S_1^{p^2}, S_2]$, $S_5^p=1$, $S_5^{-1}S_1S_5=S_2S_1$,
$S_5^{-1}S_2S_5=S_1^{\alpha^n p^2}S_2$ $(n=0, 1)$.

If G_4 is of type (4), and its generator is $S_1^{p^4}=1$, then

38. $H=[S_1^p]$, $K=[S_1^{p^3}]$, $S_5^p=1$.

For $p>3$ these groups are all conformal with abelian groups. For $p=3$ these groups are all conformal with abelian groups except groups 4, 19 $(n=1)$, 23, and 32 $(n=0, 1)$, and groups 7, 8, 25, and 25_1 which do not exist. For $p=3$ groups 4 and 5 are the same group. Groups 20_1, 32, and 33 do not exist for $p>3$.

The total number of non-abelian groups of order p^5, which contain an abelian subgroup of order p^4 is $p+39$ $(p-1\neq 3k)$, and $p+41$ $(p-1=3k)$.

In groups 7 and 8 $S_5^{-3}S_4S_5^3=S_1S_2^3S_3^3S_4=S_1S_4$, hence S_5^3 is not in G_4.

In 19 consider the isomorphism $\begin{pmatrix}S_5, & S_1, & S_2, & S_3\\ S_5^\mu, & S_1, & S_2, & S_3^\mu\end{pmatrix}$ between the group in which $S_5^{-1}S_3S_5=S_1^{xp}S_3$, and the group in which $S_5^{-1}S_3S_5=S_1^{yp}S_3$. Hence the groups are isomorphic if the congruence

$$x\equiv y\mu^2 \mod p \tag{7}$$

is satisfied. If $p=3$ then $(S_1^aS_5^b)^3=S_1^{3a(\alpha^n+1)}S_5^{3b}$, hence all the operators in 19 $(n=1)$ which are not in G_4 are of order 3, and in 20_1 all these operators are of order 9.

In 25 if $p=3$ $S_5^{-3}S_1S_5^3=S_1^3S_2^3S_3^3S_1$, hence S_5^3 is not in G_4. For $p>3$ consider the isomorphism $\begin{pmatrix}S_5, & S_1, & S_3, & S_2\\ S_5^\mu, & S_1, & S_3^\mu, & S_2^{\mu^2}\end{pmatrix}$ between the group in which $S_5^{-1}S_2S_5=S_1^{xp}S_2$ and the group in which $S_5^{-1}S_1S_5=S_1^{yp}S_2$. These groups are isomorphic if the congruence

$$x\equiv y\mu^3 \mod p \tag{8}$$

is satisfied. Hence the total number of distinct groups is equal to the common divisor of $p-1$ and 3.

In 30 consider the isomorphism $\begin{pmatrix}S_5, & S_1, & S_2\\ S_5^\mu, & S_1, & S_2^\mu\end{pmatrix}$ between the group in which $S_5^{-1}S_2S_5=S_1^{xp}S_2^{ap}S_2$ and the group in which $S_5^{-1}S_2S_5=S_1^{yp}S_2^{bp}S_2$. These groups are isomorphic if the congruences

$$x \equiv y\mu^2 \mod p, \qquad a\mu \equiv b\mu^2 \mod p \tag{9}$$

are satisfied. These congruences will not be satisfied for values of a and $b \leqq (p-1)/2$, and x and y equal to 1, or α ($a \neq b$ when $x = y$). It can be shown that those groups in which $\beta^2 + 4\alpha^n = 0$, there are $p^2(p-1)$ operators of order p^2 in G_4 which are transformed into themselves multiplied by their pth power, if $\beta^2 + 4\alpha^n$ is a quadratic residue there are $2p^2(p-1)$ operators, and if $\beta^2 + 4\alpha^n$ is a quadratic non-residue there are no operators of order p^2 in G_4 which are transformed into themselves multiplied by their pth power.*

In case $p = 3$ we have the additional groups 32 and 33, for if $S_5^{-1} S_2 S_5 = S_1 S_2$, and $S_5^{-1} S_1 S_5 = S_1^{3x} S_2^{3a} S_1$, then $S_5^{-3} S_2 S_5^3 = S_1^{3a} S_1^3 S_2$, and hence if $a = 2$ then S_5^3 is in G_4. Consider the isomorphism $\begin{pmatrix} S_5, & S_2, & S_1 \\ S_5^c S_2^d, & S_2^b, & S_1^{(\ )p} S_1^{bc(c-1)} S_2^{bc} \end{pmatrix}$ between the group in which $S_5^{-1} S_1 S_5 = S_1^{3x} S_2^6 S_1$ and the group in which $S_5^{-1} S_1 S_5 = S_1^{3y} S_2^6 S_1$. These groups are isomorphic if the congruence

$$bcx \equiv 2b(c-1) + by \mod 3 \tag{10}$$

is satisfied. Hence we may suppose x to be 0, or 2. If we suppose $S_5^3 = S_1^{3\beta_1}$ and $S_5^3 = S_1^{3\beta_2}$, then we have for the two cases the additional congruences

$$bc\beta_1 = \beta_2 c + d \mod 3, \qquad bc\beta_1 = \beta_2 c \mod 3. \tag{11}$$

This follows from the fact that in the first case $(S_5^c S_2^d)^3 = S_5^{3c} S_1^{3d}$, and in the second case $(S_5^c S_2^d)^3 = S_5^{3c}$, and hence there is but the one group in the first case, and two in the second.†

In 37 consider the isomorphism $\begin{pmatrix} S_5, & S_1, & S_2 \\ S_5^\mu, & S_1, & S_2^\mu \end{pmatrix}$, hence the congruence

$$x \equiv y\mu^2 \mod p. \tag{12}$$

5. **Groups of order p^5 which do not contain an abelian subgroup of order p^4.** It has been shown that these groups must contain an invariant abelian subgroup of order p^3, and hence must contain a subgroup of order p^4 which has a commutator subgroup of order p. The

---

* Cf. On page 198 the groups in the fourth row to be correct should read $E_3^p = E_1 E_2^{\varepsilon^{2m+1}}$.

† On page 217 consider the isomorphism $\begin{pmatrix} S_5, & S_4, & S_3, & S_2, & S_1 \\ S_5, & S_4^2, & S_3^2, & S_2^2, & S_1^2 \end{pmatrix}$. It results that the third and fourth groups are isomorphic.

groups of order p^4 which contain a commutator subgroup of order p are the groups 1, 4, 5, 6, 7, and 9 as given above. For convenience these groups may be defined by $H = [S_1, S_2]$, $K = [S_1]$, $S_4^{-1} S_3 S_4 = S_1 S_3$, and defining the pth powers of S_2, S_3, S_4, $(S_1^p = 1)$ as follows

	1	4	5	6	7	9
$S_2^p =$	1	1	1	1	S_1	S_1
$S_3^p =$	1	1	S_1	1	1	S_2
$S_4^p =$	1	S_1	S_2	S_2	1	1.

We shall designate these groups by G_4^1, G_4^4, etc.

A group of order p^5 can not contain a central of order p^2 and a commutator subgroup of order p, or p^2 without containing an abelian subgroup of order p^4.

When determining all the groups of order p^5 which contain a given subgroup of order p^4 we shall give the central H and the commutator subgroup K and the manner in which G_4 is transformed by an operator S_5 not in G_4. Since for $p > 3$ $(S_\alpha^a S_5^b)^p = S_\alpha^{ap} S_5^{bp}$, we may suppose the pth power of S_5 to be the identity, or an independent generator of G_4 which is in H.

	H	K	$S_5^{-1} S_4 S_5 S_4^{-1}$	$S_5^{-1} S_3 S_5 S_3^{-1}$	$S_5^{-1} S_2 S_5 S_2^{-1}$	S_5^p
	Groups containing G_4^1.					
39.	$[S_1]$	$[S_1]$	1	1	S_1	S_1^n $(n = 0, 1)$;
40.	$[S_1]$	$[S_1, S_2]$	S_2	1	S_1	S_1^n $(n = 0)$ $(n = 1, p > 3)$;
41.	$[S_1]$	$[S_1, S_3]$	$S_3^{\alpha^n}$	1	S_1	S_1 $(n = 0, 1)$;
42.	$[S_1]$	$[S_1, S_2, S_3]$	S_3	S_2	S_1	1 $(p > 3)$;
43.	$[S_1]$	$[S_1, S_2, S_3]$	$S_3^{\alpha^n}$	S_2 $(n = 0, 1)$, $(p - 1 = 4k)$	S_1	S_1 $(p > 3)$ $(n = 0, 1, 2, 3)$;
44.	$[S_1, S_2]$	$[S_1, S_2, S_3]$	S_3	S_2	1	1;
45.	$[S_1, S_2]$	$[S_1, S_2, S_3]$	$S_3^{\alpha^n}$	S_2	1	S_1 $(n = 0, 1)$;
46.	$[S_1, S_2]$	$[S_1, S_2, S_3]$	S_3	S_2	1	S_2 $(p > 3)$. For $p = 3$ same as 45 $(n = 1)$.
	Groups containing G_4^6 but not G_4^1.					
47.	$[S_1, S_2]$	$[S_1, S_2, S_3]$	S_3	$S_1^m S_2^{\alpha^n}$	1	S_1 $(p > 3)$ $(n = 0, 1)$ $\left(m = 0, 1, \cdots, \frac{p-1}{2}\right)$ $(p = 3)$ $(n = m = 0)$ $(n = m = 1)$;
48.	$[S_1, S_2]$	$[S_1, S_2, S_3]$	S_3^α	$S_1^m S_2^{\alpha^n}$	1	S_1 $(m^2 + 4\alpha^n \equiv 0)$.

Groups containing G_4^5 but not G_4^1, nor G_4^6.

49.	$[S_1]$	$[S_1, S_3]$	S_3^{-1}	1	S_1	S_3;
50.	$[S_1]$	$[S_1, S_2, S_3]$	S_3	$S_1^m S_2^2$	S_1	$1\,(p=3)$ $(m=1,2)$.*

Groups containing G_4^7 but not G_4^1, G_4^6, nor G_4^5.

51.	$[S_1]$	$[S_1, S_3]$	S_3	1	S_1	1.

Groups containing G_4^4 but not G_4^1, G_4^6, G_4^5, nor G_4^7.

52.	$[S_1]$	$[S_1, S_2, S_3]$	$S_3^{\alpha^n}$	S_2	S_1	$1\,(p>3)\,(n=0)$ $(p-1=3k)\,(n=0,1,2)$;
53.	$[S_1, S_2]$	$[S_1, S_2, S_3]$	S_3	S_2	1	S_2;
54.	$[S_1]$	$[S_1, S_3]$	S_3	1	S_1	$S_1\,(p=3)$.

Since $(S_\alpha S_\beta)^p = S_\alpha^p S_\beta^p$ for $p>3$, it follows that for $p>3$ all the groups of order p^5 which contain G_4^1 will contain only one such subgroup if S_5 is not of order p.

For $p=3$ group 40 ($n=0$) contains operators of order 9, and the two groups $n=0$, and 1 are isomorphic.

In 41 consider the isomorphism $\begin{pmatrix} S_5, & S_4, & S_3, & S_2, & S_1 \\ S_5^a, & S_4^b, & S_3^c, & S_2, & S_1^a \end{pmatrix}$ between the group in which $S_5^{-1} S_4 S_5 = S_3^x S_4$ and the group in which $S_5^{-1} S_4 S_5 = S_3^y S_4$. Hence the groups are isomorphic if the congruences

$$(13) \quad a \equiv bc \mod p, \qquad cx \equiv aby \mod p, \qquad x \equiv b^2 y \mod p$$

are satisfied. The groups 40 and 41 can not be isomorphic if $s^p \neq 1$, for in 40 K is the central of G_4^1 but this is not the case in 41.

It is evident that for $p=3$, H of order p, and K of order p^3, that for G_5 to exist it is necessary for $G_4 = G_4^5$.

In 43 consider the isomorphism $\begin{pmatrix} S_5, & S_4, & S_3, & S_2, & S_1 \\ S_5^a, & S_4^b, & S_1^{(\,)} S_2^{(\,)} S_3^c, & S_1^{(\,)} S_2, & S_1^a \end{pmatrix}$ between the group in which $S_5^{-1} S_4 S_5 = S_3^x S_4$ and the group in which $S_5^{-1} S_4 S_5 = S_3^y S_4$. Hence the following congruences

$$(14) \quad a \equiv bc \mod p, \quad cx \equiv aby \mod p, \quad ac \equiv 1 \mod p, \quad x \equiv a+y \mod p.$$

For $p>3$ groups 45 and 46 can not be isomorphic because 46 contains G_4^4 as a subgroup while 45 does not. In 45 if $n=1$ $(S_4^2 S_5)^3 = S_2$.

* Cf. The group for $m=2$ is omitted.

Groups containing G_4^6 will contain G_4^1 if S_5 is of order p, hence we may suppose $S_5^p = S_1$. In 47 consider the isomorphism between the group in which $S_5^{-1} S_4 S_5 = S_3^{x_3} S_4$, $S_5^{-1} S_3 S_5 = S_1^{x_1} S_2^{x_2} S_3$, and the group in which $S_5^{-1} S_4 S_5 = S_3^{y_3} S_4$, $S_5^{-1} S_3 S_5 = S_1^{y_1} S_2^{y_2} S_3$, such that S_4 corresponds to $S_4^{a_4} S_5^{a_5}$ and S_5 to $S_4^{b_4} S_5^{b_5}$ $(\mu = a_4 b_5 - a_5 b_4 \neq 0)$ and S_3 to $S_1^{(\ )} S_2^{(\ )} S_3^c$. These groups are isomorphic if the following congruences are satisfied $(p > 3)$

$$(15) \qquad \begin{aligned} & c a_4 + c a_5 y_1 \equiv b_5, \quad c a_5 y_2 \equiv b_4, \quad c x_3 \equiv \mu y_3, \\ & b_5 x_1 + a_5 x_2 \equiv c b_4 + c b_5 y_1, \quad b_4 x_1 + a_4 x_2 \equiv c b_5 y_2 \mod p. \end{aligned}$$

Combining the first two congruences $\mu = c\,(a_4^2 + a_4\, a_5\, y_1 - a_5^2\, y_2)$. Hence c and μ are independent except when $a_4^2 + a_4\, a_5\, y_1 - a_5^2\, y_2$ is a perfect square, or when $y_1^2 + 4\, y_2 \equiv 0 \bmod p$. In this case c and μ must both be a quadratic residue or both be a quadratic non-residue. Hence we may suppose $S_5^{-1} S_4 S_5 = S_3 S_4$ for 47, and $S_5^{-1} S_4 S_5 = S_3^{\alpha} S_4$ if $y_1^2 + 4 y_2 = 0$ for 48. Solving for x_1 and x_2 we have

$$x_1 \equiv c y_1 \mod p, \qquad x_2 \equiv c^2 y_2 \mod p.$$

These groups are characterized by the fact that if $m^2 + 4\alpha^n \equiv 0$ the groups contain but one G_4^4, if $m^2 + 4\alpha^n$ is a quadratic residue there are two such subgroups, and if $m^2 + 4\alpha^n$ is a quadratic non-residue G_5 does not contain G_4^4.* For $p = 3$, $(S_4^a\, S_5^b)^3 = S_1^A\, S_2^B$ where $A = b + a b^2 y_1\, y_3 + 2\, b a^2\, y_3$, and $B = a + a b^2\, y_2\, y_3$. It is not difficult to show that there are four sets of isomorphic groups

$$S_5^{-1} S_4 S_5 S_4^{-1} = \begin{pmatrix} S_3, & S_3^2 \\ S_2, & S_2 \end{pmatrix}, \quad \begin{pmatrix} S_3, & S_3 \\ S_1 S_2^2, & S_1^2 S_2^2 \end{pmatrix}, \quad \begin{pmatrix} S_3 & S_3, & S_3^2, & S_3^2 \\ S_1 S_2, & S_1^2 S_2, & S_1 S_2^2, & S_1^2 S_2^2 \end{pmatrix},$$

$$S_5^{-1} S_3 S_5 S_3^{-1} = \begin{pmatrix} S_3, & S_3^2, & S_3^2 \\ S_2^2, & S_1 S_2, & S_1^2 S_2 \end{pmatrix}.$$

The groups in the fourth set are isomorphic to groups containing G_4^1. The groups in the first two sets do not contain G_4^4 as subgroup, while those in the third contain one subgroup G_4^4.

In 49 $S_5^{-p} S_4 S_5^p = S_1 S_4$, therefore $S_5^p = S_3^{-1}$. This is the only group which contains G_4^9.

In 50 if we suppose $S_5^{-1} S_3 S_5 = S_1^x S_2^y S_3$, then $S_5^{-3} S_4 S_5^3 = S_1^y S_3^3 S_4$. Hence, if $y = 2$, S_5^3 is commutative with S_4, but for $(p > 3)$ S_5^p can not be commutative with S_4 and therefore the groups do not exist for $p > 3$. Since $(S_4^a S_5^b)^3 = S_1^{a x + a + 2 a^2 b}$, then if $x = 2$ the operator $S_4^a S_5^b$ $(a \neq 0)$ is of order 9, but if $x = 0$, or 1 then this operator is of order 3 for $a = b$, or $2\,b$ respectively. It is not difficult to verify that the two groups in which $x = 0$, and $x = 1$ are isomorphic.

* Cf. On page 183 the groups in the third row to be correct should read $E_4^p = E_1\, E_2^{\varepsilon^{2m+1}}$.

In 52 consider the isomorphism $\begin{pmatrix} S_5, & S_4, & S_3, & S_2, & S_1 \\ S_5^a, & S_4^b, & S_1^{(\,)}\, S_2^{(\,)}\, S_3, & S_1^{(\,)}\, S_2^a, & S_1^b \end{pmatrix}$ between the group in which $S_5^{-1} S_4 S_5 = S_3^x S_4$ and the group in which $S_5^{-1} S_4 S_5 = S_3^y S_4$. Hence the congruences

$$a^2 \equiv b \mod p, \qquad x \equiv a^3 y \mod p. \tag{16}$$

Group 53 contains $p+1$ subgroups G_4 $(p>3)$, but only 2 such subgroups for $p=3$.

In 54 since $(S_4^a S_5^b)^3 = S_1^{2a^3b+b} S_4^{3a}$, the group 41 in which S_4 is of order three is distinct from the group in which S_4 is of order nine.

All these groups in which $p>3$ are conformal with abelian groups. All the groups in which $p=3$ are conformal with abelian groups except groups 40, 44, and 50.

The total number of non-abelian groups of order p^5 which do not contain an abelian subgroup of order p^4 is $p+19$ $(p=12k-1)$, $p+21$ $(p=12k\pm5)$, or $p+23$ $(p=12k+1)$, and 17 $(p=3)$.

REPRESENTATION OF FUNCTIONS DETERMINED BY THEIR INITIAL VALUES.*

By W. J. Trjitzinsky.

Among the functions of a real variable which are indefinitely derivable there are found extensive classes of functions which are determined by their initial values, that is by the values of the function and of its successive derivatives at, say, $x = 0$. Among functions determined by initial values there are functions either analytic, or quasi-analytic (their Taylor's expansions diverge) or neither analytic nor quasi-analytic.† De la Vallée-Poussin was the first to give methods for constructing functions (analytic, quasi-analytic or neither) assuming an assigned set of initial values for $x = 0$. One of his theorems is:‡

"Let $m_1, m_2, \cdots, m_n, \cdots$ be a sequence of positive integers such that for all n,

$$\frac{m_{n+1}}{m_n} > r > 1.$$

Let $F_0, F_2, \cdots, F_{2k}, \cdots$ be a sequence of bounded numbers. It is always possible to determine the coefficients of the trigonometric series

$$F(x) = \alpha_1 \cos m_1 x + \alpha_2 \cos m_2 x + \cdots + \alpha_n \cos m_n x + \cdots$$

so that this series and all its derivatives will converge, and the sum function and its derivatives of even order will take on respectively the given values $F_0, F_2, \cdots$ for $x = 0$."

Examining the proof of this theorem it is observed that if the trigonometric series $F(x) = \sum_{i=1}^{\infty} \alpha_i \cos m_i x$ be replaced by $F(x) = \sum_{i=1}^{\infty} \alpha_i f(m_i x)$ where $f(x) = b_0 + b_2 x^2 + b_4 x^4 + \cdots$ with $b_{2j} \neq 0$ $(j = 0, 1, 2, \cdots)$ and such that $|f(x)|$ and $|f^{(2p)}(x)|$ are $< M$ for all p and all real x, the theorem still holds true provided the ratios $F_{2n}/f^{(2n)}(0)$ remain bounded for all n.

For in the course of his proof Vallée-Poussin shows that if $C_0, \cdots, C_{2n-2}, \cdots$ are a bounded set of constants, and if the a_k satisfy the equations

* Received May 14, revised June 16, 1927.

† Borel introduced quasi-analytic functions of a complex variable in "Les séries de fonctions analytiques et les fonctions quasi-analytiques", Comptes Rendus, vol. 154 (1912), p. 1491. Denjoy has developed the subject for functions of a real variable in Comptes Rendus, vol. 173 (1921), p. 1329.

‡ The Rice Institute Pamphlet, vol. 12 (1925), no. 2, p. 165.

$$(-1)^p C_{2p} = \sum_{k=1}^{n} m_k^{2p} a_k, \quad p = 0, 1, 2, \cdots, (n-1), \tag{1}$$

then $\lim_{n\to\infty} a_k$ exists and is equal to α_k which is given by the formula

$$\alpha_k = \frac{1}{2\pi i} \int_C \frac{\varphi(z^2)\,\psi(z)\,dz}{(z^2 - m_k^2)\,\varphi^{(1)}(m_k^2)}, \tag{2}$$

$$\psi(z) = \sum_{n=0}^{\infty} (-1)^n C_{2n} \frac{1}{z^{2n+1}}, \quad \varphi(z) = \prod_{j=1}^{\infty} \left(1 - \frac{z}{m_j^2}\right).*$$

Further

$$|a_k| \text{ and } |\alpha_k| \text{ are } < h' \left(\frac{m_1 m_2 \cdots m_{k-1}}{m_k^{k-1}}\right)^2 \tag{3}$$

for all k, h' being a constant depending on the C_{2p}.

If now

$$F(x) = \sum_{k=1}^{\infty} \alpha_k f(m_k x),$$

then obviously we can determine the α_k as the limits of the a_k if in equations (1) and (2), C_{2p} will be replaced by $(-1)^p F_{2p}/f^{(2p)}(0)$. The inequalities (3) together with the boundedness of $|f(x)|$ and $|f^{(2p)}(x)|$ for all p and real x then secure the convergence of the series for $F(x)$. The result may be stated in the following form.

THEOREM I. *Let $m_1, m_2, \cdots, m_n, \cdots$ be a sequence of positive integers such that for all n*

$$\frac{m_{n+1}}{m_n} > r > 1;$$

let $f(x) = b_0 + b_2 x^2 + \cdots + b_{2n} x^{2n} + \cdots$ (all powers present) and $|f(x)|$ and $|f^{(2p)}(x)|$ be $< M$ for all p and all real x, and let $F_0, F_2, \cdots, F_{2k}, \cdots$ be a sequence of constants such that $F_{2k}/f^{(2k)}(0)$, $(k = 0, 1, 2, \cdots)$ are bounded. It is always possible to determine the coefficients of the series

$$F(x) = \alpha_1 f(m_1 x) + \alpha_2 f(m_2 x) + \cdots + \alpha_n f(m_n x) + \cdots$$

so that the series and all its derivatives will converge, and the function and its derivatives of even order will take on the given values $F_0, F_2, \cdots, F_{2k}, \cdots$ for $x = 0$. (α_k will be given by (2) where $C_{2p} = (-1)^p F_{2p}/f^{(2p)}(0)$).

The sum function $F(x)$ is indefinitely derivable. Let us examine, however, conditions under which $F(x)$ will be analytic.

---

* C denotes a circle with radius R, such that $R > |F_{2p}/f^{(2p)}(0)|$ for all p, having origin for center.

$$|F^{(2p)}(x)| = |a_1 m_1^{2p} f^{(2p)}(m_1 x) + a_2 m_2^{2p} f^{(2p)}(m_2 x) + \cdots + a_n m_n^{2p} f^{(2p)}(m_n x) + \cdots|$$
$$< M\{m_1^{2p}|a_1| + m_2^{2p}|a_2| + \cdots + m_n^{2p}|a_n| + \cdots\}.$$

Using inequalities (3)

$$|F^{(2p)}(x)| < M\{L_p + R_p\}$$

where

$$L_p = h' \cdot [m_1^{2p} + (m_2^{2p-2} \cdot m_1^2) + (m_3^{2p-4} \cdot m_1^2 \cdot m_2^2) + \cdots + (m_{p-1}^4 m_1^2 \cdots m_{p-2}^2) + 2(m_1^2 \cdots m_p^2)]$$
$$< h'(m_1 m_2 \cdots m_p)^2 (p+1)$$

and

$$R_p = h'\left[\frac{m_{p+1}^2}{m_{p+2}^2} + \frac{m_{p+1}^2 m_{p+2}^2}{m_{p+3}^4} + \cdots + \frac{m_{p+1}^2 \cdots m_{p+k}^2}{m_{p+k+1}^{2k}} + \cdots\right](m_1 m_2 \cdots m_p)^2$$

but $m_j/m_k < \frac{1}{r} < 1$ if $j < k$, since $m_{j-1}/m_j < \frac{1}{r}$ for all j by hypothesis; hence

$$R_p < \frac{h'(m_1^2 m_2^2 \cdots m_p^2)}{r^2 - 1}$$

and

$$|F^{(2p)}(x)| < Mh'(m_1 m_2 \cdots m_p)^2 \left[p + \frac{r^2}{r^2-1}\right]. \tag{4}$$

The second member of (4) is independent of x and increases with p. It is known* that if $|\sqrt[2p]{F^{(2p)}(x)}| < kp$ for all p, $F(x)$ is analytic on the segment (0,1).

Hence, if

$$\varphi(p) = Mh'(m_1 m_2 \cdots m_p)^2 \left[p + \frac{r^2}{r^2-1}\right] < (kp)^{2p} \tag{5}$$

for some fixed k and all p, the sum function $F(x)$ is an analytic function on (0,1) (at least). For $x = 0$, it and its derivatives of even order assume a set of values $F_0, F_2, \cdots, F_{2n}, \cdots$, and therefore on (0,1)

$$F_0 + \frac{F_2}{2!}x^2 + \frac{F_4}{4!}x^4 + \cdots + \frac{F_{2n}}{(2n)!}x^{2n} + \cdots = \alpha_1 f(m_1 x) + \alpha_2 f(m_2 x) + \cdots.$$

When $m_n = n^n$, the inequality (5) as well as inequalities corresponding to different Denjoy classes of quasi-analytic functions

* The Rice Institute Pamphlet, vol. 12, (1925), no. 2, p. 147.

$$\varphi(p) < [kp \log(2p) \log_2(2p) \cdots \log_j(2p)]^{2p} \qquad (j = 1, 2, \ldots)$$

do not hold. In this case our method does not establish analyticity or quasi-analyticity of $F(x)$. It is interesting to note that with $m_n = n^n$, in Vallée-Poussin's case, that is when $f(x) = \cos x$, $F(x)$ though being determined by the initial values is neither analytic nor quasi-analytic.*

We sum up in form of a theorem.

THEOREM II. *If $f(x)$ satisfies conditions of Theorem* I *and $F(x)$ is an assigned even function analytic in the neighborhood of the origin and such that $F^{(2n)}(0)/f^{(2n)}(0)$ are bounded for all n, then*

$$F(x) = \alpha_1 f(m_1 x) + \alpha_2 f(m_2 x) + \cdots + \alpha_n f(m_n x) + \cdots$$
$$F^{(j)}(x) = \alpha_1 m_1^j f^{(j)}(m_1 x) + \alpha_2 m_2^j f^{(j)}(m_2 x) + \cdots + \alpha_n m_n^j f^{(j)}(m_n x) + \cdots \qquad (j = 1, 2, \ldots)$$

at least on the segment $(-1, +1)$ provided that m_i satisfy the inequality (5) *and conditions of Theorem* I. *Coefficients α_k of this expansion are given by* (2) *where $c_{2p} = (-1)^p F_{2p}/f^{(2p)}(0)$.*

COROLLARY. If $f(x)$ and $F(x)$ satisfy conditions of Theorem II, then

$$F(x) = \alpha_1 f(rx) + \alpha_2 f(r^2 x) + \alpha_3 f(r^3 x) + \cdots + \alpha_n f(r^n x) + \cdots \quad (r > 1)$$

and

$$F^{(j)}(x) = \alpha_1 r^j f^{(j)}(rx) + \alpha_2 r^{2j} f^{(j)}(r^2 x) + \cdots + \alpha_n r^{nj} f^{(j)}(r^n x) + \cdots \qquad (j = 1, 2, \ldots)$$

at least on $(-1, +1)$.

This follows from the fact that $m_j = r^j$ satisfy (5).

In connection with the results of Theorems I and II it is natural to investigate:

(I) The order of approximation furnished by the expansion

$$F(x) = \sum \alpha_i f(m_i x);$$

(II) Construction of functions $f(x)$ which can be used in this expansion, and determination of some of their properties.

We have

$$\alpha_k < h_1 \left(\frac{m_1 m_2 \cdots m_{k-1}}{m_k^{k-1}} \right)^2.$$

* The Rice Institute Pamphlet, vol. 12 (1925), no. 2, pp. 171–2.

Let us form:

$$R_n(x) = |F(x) - [a_1 f(m_1 x) + a_2 f(m_2 x) + \cdots + a_n f(m_n x)]|$$
$$= |a_{n+1} f(m_{n+1} x) + a_{n+1} f(m_{n+2} x) + \cdots + a_{n+k} f(m_{n+k} x) + \cdots|$$
$$< h_1 M \left| \left(\frac{m_1 \cdots m_n}{m_{n+1}^n}\right)^2 + \left(\frac{m_1 m_2 \cdots m_{n+1}}{m_{n+2}^{n-1}}\right)^2 + \cdots \right|$$
$$< h_1 M \left\{ \frac{1}{r^{2n}} + \frac{1}{r^{2n+2}} + \cdots + \frac{1}{r^{2n+2j}} + \cdots \right\}$$
$$= h_1 M \frac{1}{r^{2n}\left(1 - \frac{1}{r^2}\right)} < \frac{L}{r^{2n}} \qquad (r > 1).$$

Hence, we have

THEOREM III. *The approximation furnished by taking the first n terms of the expansion defined by Theorems* I *or* II *is of order of* $\frac{1}{r^{2n}}$ $(r > 1)$.

Studying the function $f(x)$, it is observed that the condition that all even powers should be present was introduced in order to have $F_{2p}/f^{(2p)}(0)$ bounded for all p. However, it is possible to have some of the powers of $f(x)$ missing if the corresponding F_{2p} are missing; let us suppose that this is the case.

Definition. $f(x) = b_0 + b_1 x^2 + b_2 x^4 + \cdots$ will be said to belong to the class (A), if for all p and all real x $f(x)$ and $f^{(2p)}(x) < M$.

It is observed that if $\varphi(x)$ is of class (A), then

$$[c_0 + c_1 \varphi(p_1 x) + c_2 \varphi(p_2 x) + \cdots]$$

and

$$[c_1 p_1^{2p} \varphi^{(2p)}(p_1 x) + c_2 p_2^{2p} \varphi^{(2p)}(p_2 x) + \cdots] < M$$

for all p and all real x, if c_i and p_j satisfy the condition

(B) $\sum_{i=0}^{\infty} |c_i|$ converges and $|p_j| \leqq 1$ $(j = 1, 2, \cdots)$, $p_j \neq p_k$.

Accordingly we have, that if $\varphi(x)$ is of class (A) and c_i, p_j satisfy (B), then $f(x) = \sum_{j=0}^{\infty} c_j \varphi(p_j x)$ is of class (A).

Vallée-Poussin* uses $\cos x$ which is of class (A); consequently, letting $\varphi(x) = \cos x$, functions of class (A) can be given by

$$f(x) = c_0 + c_1 \cos p_1 x + c_2 \cos p_2 x + \cdots \tag{6}$$

where c_i, p_j satisfy (B); (6) will be periodic only in particular cases.

* The Rice Institute Pamphlet, vol. 12 (1925), no. 2, p. 165.

Differentiating (6) in succession and letting $x = 0$ after each differentiation, b_i in $f(x) = \sum_{i=0}^{\infty} b_i x^{2i}$ are expressed in terms of c_j, p_k.

We state the result in the form of a theorem.

THEOREM IV. *If* b_n $(n = 0, 1, 2, \cdots)$ *are determined by the expressions* $b_n = (-1)^n [p_1^{2n} c_1 + p_2^{2n} c_2 + \cdots + p_k^{2n} c_k + \cdots]$ *where* c_i, p_j *satisfy* (B), *then* $f(x) = \sum_{j=0}^{\infty} b_j x^{2j}$ *is of class* (A).

Applying Theorem II we have

THEOREM V. *Any function* $f(x)$ *of class* (A) *satisfies at least on* $(-1, +1)$ *the relation*

$$f(x) = \alpha_1 f(m_1 x) + \alpha_2 f(m_2 x) + \cdots + \alpha_k f(m_k x) + \cdots$$

where m_i *satisfy conditions of Theorem* II; *coefficients* α_i *are independent of the choice of* $f(x)$; *they are given by*

$$\alpha_i = \frac{1}{2\pi i} \int_C \frac{z\,\varphi(z^2)\,dz}{(z^2-1)(z^2-m_k^2)\,\varphi^1(m_k^2)},$$

where $\varphi(z) = \prod_{j=1}^{\infty} \left(1 - \frac{z}{m_j^2}\right)$, *and* C *is a circle with origin for center and radius* $R > 1$.

AN INTEGRATION METHOD OF SUMMING SERIES.*

By Glenn James.

1. **General formula.** We consider only series for which the general term, a_i, looked upon as a function of the subscript is integrable on the interval, $(1,\ n+1)$, n being an arbitrary positive integer. From the calculus

$$\int_1^{n+1} a_i\,dx - \sum_1^n \int_i^{i+1} a_x\,dx = 0.$$

Omitting the differential and adding $\sum_1^n a_i$ to both members of this identity changes it into

$$\sum_1^n a_i = \int_1^{n+1} a_x + \sum_1^n \left[a_x - \int_i^{i+1} a_x\right]. \tag{1}$$

We now apply (1) to the series, $\sum_1^n \left[a_i - \int_i^{i+1} a_x\right]$, and substitute the result in (1). This gives

$$\sum_1^n a_x = 2\int_1^{n+1} a_x - \int_1^{n+1}\int_x^{x+1} a_x + \sum_1^n \left[a_i - 2\int_i^{i+1} a_x + \int_i^{i+1}\int_x^{x+1} a_x\right]. \tag{2}$$

Applying (1) to the summation on the right, as before, we obtain

$$\begin{aligned}\sum_1^n a_i = {} & 3\int_1^{n+1} a_x - 3\int_1^{n+1}\int_x^{x+1} a_x + \int_1^{n+1}\int_x^{x+1}\int_x^{x+1} a_x \\ & + \sum_1^n \left[a_i - 3\int_i^{i+1} a_x + 3\int_i^{i+1}\int_x^{x+1} a_x - \int_i^{i+1}\int_x^{x+1}\int_x^{x+1} a_x\right].\end{aligned} \tag{3}$$

which can be written in the form

$$\sum_1^n a_i = \int_1^{n+1} f_3(a_x) + \sum_1^n \left[a_i - \int_i^{i+1} f_3(a_x)\right]$$

where

$$f_3(a_x) \equiv 3a_x - 3\int_x^{x+1} a_x + \int_x^{x+1}\int_x^{x+1} a_x.$$

In general

$$\sum_1^n a_i = \int_1^{n+1} f_m(a_x) + \sum_1^n \left[a_i - \int_x^{x+1} f_m(a_x)\right]. \tag{4}$$

* Received June 14, 1927.

where*

$$f_m(a_x) \equiv \sum_{r=2}^{m+1} (-1)^r \binom{m}{r} \left(\int_x^{x+1}\right)^{r-2} a_x,$$

the form $\left(\int_x^{x+1}\right)^{r-2} a_x$ denoting the $(r-2)$tuple integral of a_x. The series, $\sum_1^n \left[a_i - \int_i^{i+1} f_m(a_x)\right]$, will be denoted by $\sum_{i=1}^n R_m(a_i)$ and spoken of as the mth remainder series. The usefulness of our summation formula depends, of course, upon the nature of this remainder series. Three cases of especial interest present themselves: that in which $\sum_1^n R_m(a_i)$ vanishes for some m and all n's, that in which it is a function of $\sum_1^n a_i$ and that in which $\lim_{m\to\infty} \lim_{n\to\infty} \sum_1^n R_m(a_i) = 0$. In the first two cases, this method is logically equivalent to the algebraic "method of differences". However the use of the integrals facilitates its application, as the following examples show.

Example (1). An important class of series which fall under case one is that class whose general term is a polynomial in i, the number of the term. The essential step in summing such series is summing $\sum_1^n i^p$. Hence we first show that $\sum_1^n R_m(a_i)$ is zero for m sufficiently large. Starting with $m = 1$, we find

$$(5) \quad \sum_1^n R_m(i^p) = \sum_1^n \left[i^p - \int_i^{i+1} x^p\right] = \sum_1^n [c_1 i^{p-1} + c_2 i^{p-2} + \cdots + c_p]$$

where c_i, $i = 1, 2, 3, \cdots, p$, are constants. Now if a_i and b_i are the general terms of two series to which our summation formula applies, we have the very obvious relations

$$(6) \quad \sum_1^n R_m(a_i + b_i) = \sum_1^n R_m(a_i) + \sum_1^n R_m(b_i),$$

$$(7) \quad \sum_1^n R_m(c\,a_i) = c \sum_1^n R_m(a_i).$$

We also have

$$(8) \quad \sum_1^n R_1(c) = 0.$$

* Obviously it is not at all necessary for $f_m(a_x)$ to have this particular form in order to make (4) valid. But this form is best suited to the purposes of this paper. See, however, Example (3).

It follows from (5), (6), (7) and (8) that

$$\sum_1^n R_{p+1}(i^p) = 0.$$

The sum, then, of $\sum_1^n i^p$ is $\int_1^{n+1} f_{p+1}(i^p)$.

Evaluating this integral and writing the result in systematic form we have*

$$\begin{aligned}\sum_1^n i^p = (n+1)^{p+1} - 1 - \frac{p}{2!(p+2)}\{[(n+2)^{p+2} - 2^{p+2}] - [(n+1)^{p+2} - 1]\} \\ + \frac{p(p-1)}{3!(p+2)(p+3)}\{[(n+3)^{p+3} - 3^{p+3}] - 2[(n+2)^{p+3} - 2^{p+3}] \\ + [(n+1)^{p+3} - 1]\} - \cdots + (-1)^p \frac{p!}{(2p+1)!}\Big\{[(n+p+1)^{2p+1} \\ - (p+1)^{2p+1}] - p[(n+p)^{2p+1} - p^{2p+1}] + \frac{p(p-1)}{2!}[(n+p-1)^{2p+1} \\ - (p-1)^{2p+1}] - \cdots + (-1)^p[(n+1)^{2p+1} - 1]\Big\}. \end{aligned} \tag{9}$$

Example (2). The ordinary Geometric Progression comes under case two. By formula (1)

$$\begin{aligned}\sum_1^n a r^{i-1} &= a(r^n - 1)\ \log\ r + \sum_1^n a[r^{i-1} - (r^i - r^{i-1})\ \log\ r] \\ &= a(r^n - 1)\ \log\ r + [1 - (r-1)\ \log\ r] \sum_1^n a r^{i-1}.\end{aligned}$$

Solving the latter equation for $\sum_1^n a r^{i-1}$ gives the familiar formula for the sum of a G. P.

Example (3). In order to sum $\sum_1^n \sin\ i\theta$ we seek a function, $f(i\theta)$, such that

$$\sin\ i\theta - \int_i^{i+1} f(x\theta) = k_1\ \sin i\theta$$

or simply

$$\int_i^{i+1} f(x\theta) = k_2\ \sin i\theta.$$

Differentiating both members gives

$$f(i\theta + \theta) - f(i\theta) = k_3\ \cos\ i\theta.$$

* This form for the sum of this series seems to be new, For the sum of this series in terms of Bernoullian Numbers see Bromwich, "Theory of Infinite Series", p. 235. See also P. Pepin, Nouv. Ann. Math., vol. 15 (1856), p. 27, and J. J. Schwatt, "Introduction to the Operations with Series", p. 81.

A solution of this functional equation, seen by inspection, is $\sin\left(i\theta-\frac{\theta}{2}\right)$. We now substitute $\sin\left(i\theta-\frac{\theta}{2}\right)$ for $f_m(a_x)$ in formula (4) and easily obtain the well-known formula for our series.

A combination of the methods used in cases one and two is effective for summing certain series. For example to sum $\sum_1^n i^p r^i$ we apply formula (1) and solve the resulting equation for $\sum_1^n i^p r^i$ as in case two, then apply (1) to the remainder series which can be broken up into series of the form $\sum_1^n i^q r^i$ where q is a positive integer less than p. The pth repetition of this dual process gives a geometric progression for the remainder.

2. **Rapid evaluation of series.** We are concerned here with those series for which $\lim_{m\to\infty}\lim_{n\to\infty}\sum_1^n R_m(a_i) = 0$. Two theorems which provide sufficient criteria for selecting such series will now be proved.

THEOREM 1. *If $\varphi_0(x)$, x greater than arbitrary X, is positive or zero and possesses derivatives to the nth order, those of odd order being negative or zero and those of even order being positive or zero and if $\varphi_i(x) \equiv \varphi_{i-1}(x) - \int_i^{i+1} \varphi_{x-1}(x)$, $i = 1, 2, 3, \cdots, n$, then $\varphi_i(x)$ and their first $n-i$ derivatives obey the same rule of signs.*

Proof: Differentiate the identity

$$\varphi_1(x) = \varphi_0(x) - \int_x^{x+1} \varphi_0(x) \tag{10}$$

obtaining

$$\varphi_1'(x) = \varphi_0'(x) - [\varphi_0(x+1) - \varphi_0(x)]. \tag{11}$$

By the mean value theorem

$$\varphi_0(x+1) - \varphi_0(x) = \varphi_0'(x+\varepsilon), \qquad |\varepsilon| < 1.$$

But

$$|\varphi_0'(x+\varepsilon)| \leqq |\varphi_0'(x)|$$

since $\varphi_0'(x)$ is negative and monotonic increasing. Consequently $\varphi_1'(x)$ is negative. Differentiating (11) gives

$$\varphi_1''(x) = \varphi_0''(x) - [\varphi_0'(x+1) - \varphi_0'(x)].$$

Again by the mean value theorem.

$$\varphi_0'(x+1) - \varphi_0'(x) = \varphi_0''(x+\varepsilon), \qquad |\varepsilon| < 1.$$

But $\varphi_0''(x+\varepsilon) \leqq \varphi_0''(x)$ since $\varphi_0'''(x)$ is negative. Hence $\varphi_1''(x)$ is positive. It follows from these arguments that $\varphi_1^{[m]}(x)$, $m \leqq n-1$, is positive or negative according as m is even or odd. Similarly it can be shown that $\varphi_2^{[m]}(x)$, $m \leqq n-2$ obeys the same rules of signs and so in general does

$$\varphi_i^{[m]}(x), \qquad m \leqq n-i, \quad i = 1, 2, 3, \cdots, n-1.$$

Finally $\varphi_i(x)$, $i \leqq n$, is positive or zero since $\int_x^{x+1} \varphi_{i-1}(x)$ lies between the maximum and minimum values of its integrand, which values are $\varphi_{i-1}(x)$ and zero.

THEOREM II. *If $\varphi_0(x)$ is positive or zero and its odd and even derivatives are respectively positive or zero and negative or zero, then $\varphi_i(x)$, $i \geqq 1$, is negative and its odd and even derivatives, up to the order $n-i$ are respectively positive or zero and negative or zero.*

Proof: The minimum value of $\varphi_0(x)$ in the interval, $(x, x+1)$, is $\varphi_0(x)$ hence

$$\int_x^{x+1} \varphi_0(x) \geqq \varphi_0(x).$$

Consequently

$$\varphi_1(x) = \varphi_0(x) - \int_x^{x+1} \varphi_0(x) \leqq 0. \tag{12}$$

Also

$$\varphi_1'(x) = \varphi_0'(x) - \varphi_0'(x+\varepsilon) \geqq 0, \qquad |\varepsilon| \leqq 1, \tag{13}$$

since $\varphi_0'(x)$ is positive and monotonic decreasing; and

$$\varphi_1''(x) = \varphi_0''(x) - \varphi_0''(x+\varepsilon_2) \leqq 0 \qquad |\varepsilon_2| \leqq 1, \tag{14}$$

because $\varphi_0''(x)$ is negative and monotonic increasing. The derivations of (13) and (14) are dependent upon the orders of the derivatives being odd and even, respectively, and not upon the specific orders of the derivatives. Consequently, we can now say that $[-\varphi_1(x)]$ is a function which satisfies the hypothesis of Theorem I. From this fact, it follows that the derivatives of $\varphi_i(x)$, $i = 2, 3, \cdots, n$, up to the order $n-i$ are positive or negative according as they are odd or even, which completes the proof of the present theorem.

When $\varphi_0(x)$ is *negative* and its odd and even derivatives have opposite signs one or the other of the above theorems applies* to $[-\varphi_0(x)]$. We are now able to prove the following useful theorem.

* If the derivatives of a function, $\varphi_0(x)$, all have the same sign, neither the odd nor the even derivatives of $\varphi_i(x)$ have permanence of sign.

THEOREM III. *If a_x, x equal to or greater than arbitrary N, has a permanent sign and has infinitely many derivatives with permanent signs, those of odd and those of even orders having opposite signs then**

$$\sum_N^n a_i = \int_N^n f_m(a_x)\,dx + e_m, \quad m > 1 \tag{15}$$

where

$$f_m(a_x) = \sum_{r=2}^{m+1} (-1)^r \binom{m}{r} \left(\int_x^{x+1}\right)^{r-2} a_x$$

$$|e_m| \leqq a_N - \int_N^{N+1} f_{m-1}(a_x) \text{ and } \lim_{m\to\infty} e_m = 0, \text{ for all } n.$$

Proof: If a_x be positive the hypotheses of this theorem are identical either with those of Theorem I or with those of Theorem II. In the first case every remainder series monotonically decreases and has all of its terms positive while in the second case every one monotonically increases but has all of its terms negative. In either case each remainder series, after the first, is numerically less than the first term of the preceding one; and, obviously, the sequence composed of the first terms of the remainder series, namely $R_1(a_i)$, $R_2(a_i)$, $R_3(a_i)$, $\cdots$, has the limit zero. The cases in which a_x is negative are merely the reflection in the x axis of the above cases.

From geometrical considerations it is easy to see that each remainder series after the first is not only numerically less than or equal to the first terms of the preceding remainder series but is numerically equal to or greater than one half of that term. Consequently, in numerical work, it is better to use for our sum formula,

$$\sum_N^n a_i = \int_N^{n+1} f_m(a_x) + \frac{3}{4}\left[a_N - \int_N^{N+1} f_{m-1}(a_x)\right] + e'_m, \quad m > 1 \tag{16}$$

where

$$|e'_m| \leqq \frac{1}{4}\left[a_N - \int_N^{N+1} f_{m-1}(a_x)\right].$$

A numerical example: Suppose we wish to evalute the series $\sum_1^\infty (-1)^i i^{-\frac{1}{3}}$. We add up the first four terms since this is not especially laborious and

* It should be noted, in the interest of the use of our formula when $m = 1$, that when a_x and its first derivative have like signs the first remainder series does not necessarily converge. However the second remainder series bears the same relation to the first that the first remainder series does to the original series in the familiar case in which a_x and its first derivative have opposite signs.

the error made in these terms by our summation formula is relatively large. We find

$$\sum_1^4 (-1)^i i^{-\frac{1}{3}} = .326940.$$

In order to apply our sum formula, it is best to group the terms and evaluate

$$\sum_5^\infty \left[(2i-1)^{-\frac{1}{3}} - (2i)^{-\frac{1}{3}}\right].$$

By the use of (16) with $m = 1$, which is permissable here since this series monotonically decreases, we find

$$\begin{aligned}\sum_5^\infty \left[(2i-1)^{-\frac{1}{3}} - (2i)^{-\frac{1}{3}}\right] &= \int_5^\infty \left[(2x-1)^{-\frac{1}{3}} - (2x)^{-\frac{1}{3}}\right] \\ &\qquad + \frac{3}{4}\left[9^{-\frac{1}{3}} - 10^{-\frac{1}{3}}\right] + e_1', \\ &= .236130 + .01244 + e_1' \\ &= .24857 + e_1'.\end{aligned}$$

Then

$$\sum_1^\infty (-1)^i i^{-\frac{1}{3}} = .326940 + .24857 + e_1' = .57551 + e_1'$$

where $|e'| \leqq .0031$. With $m = 2$, in (16), we have for the sum after the 8th term

$$\begin{aligned}&2\int_5^\infty \left[(2x-1)^{-\frac{1}{3}} - (2x)^{-\frac{1}{3}}\right] - \int_5^\infty \int_x^{x+1} \left[(2x-1)^{-\frac{1}{3}} - (2x)^{-\frac{1}{3}}\right] \\ &\quad + \frac{3}{4}\left[9^{-\frac{1}{3}} - 10^{-\frac{1}{3}}\right] - \int_5^6 \left[(2x-1)^{-\frac{1}{3}} - (2x)^{-\frac{1}{3}}\right] + e_2',\end{aligned}$$

which evaluates into $.24524 + e_2'$. The sum of the original series is then

$$.572179 + e_2'$$

where $|e_2'| \leqq .0005$.

By ordinary arithmetic computation, over a billion terms would have to be used to reach this degree of accuracy.*

The sum to M terms of $\sum i^{-r}$, $r \geqq 0$. Adding the first ten terms and then applying (15), we have, for $m = 1$,

$$\begin{aligned}\sum_1^n i^{-r} &= \sum_1^{10} i^{-r} + \int_{11}^{n+1} x^{-r}\,dx + e_1, \\ &= \sum_1^{10} i^{-r} + \frac{(n+1)^{1-r}}{1-r} - \frac{(11)^{1-r}}{1-r} + e_1\end{aligned}$$

* Compare with Euler's method of evaluating slowly convergent series. Bromwich, "Theory of Infinite Series", p. 55.

where $|e_1| \leqq 1$, since zero makes 11^{-r} a maximum. For $m = 2$,

$$(17) \qquad \sum_1^n i^{-r} = \sum_1^{10} i^{-r} + 2\int_{11}^{n+1} x^{-r}\,dx - \int_{11}^{n+1}\int_x^{x+1} x^{-r}\,dx + e_2$$

where

$$(18) \qquad |e_2| \leqq \left[11^{-r} - \frac{(12)^{1-r}}{1-r} + \frac{(11)^{1-r}}{1-r}\right].$$

In order to make (17) a worthwhile formula, for each r with all n, we must determine an upper limit of $|e_2|$. Such an upper limit exists since the right member of (18) is a continuous function of r and is zero for r equal to zero and for r infinite. Considering only the first two decimal places, in non-integral values of r, the desired upper limit of $|e_2|$, that is the greatest value of the right member of (18), is found to be .00669, occurring for $r = .4$. Proceeding, as above, with larger values of m one can secure a formula for the sum to n terms of $\sum i^{-r}$, $r > 0$, which is correct to any desired degree of accuracy independent of r and n. The limits of the errors in (17) and similar formulas, given by other values of m, can be reduced to one fourth of the above values by adding to the sum formulas three fourths of the difference between the first and the nth term of the preceding Remainder Series. For we have

$$R_m(11^{-r}) - R_m(n^{-r}) \geqq e_m \geqq \frac{1}{2}[R_m(11^{-r}) - R_m(n^{-r})].$$

Subtracting three fourths of the first member of this inequality from each member of the inequality gives

$$|e_m| \leqq \frac{1}{4}[R_m(11^{-r}) - R_m(n^{-r})],$$

whence, for n infinite as in (16),

$$|e'_m| \leqq \frac{1}{4} R_m(11^{-r}).$$

We can now write*

$$(19) \qquad \begin{aligned} \sum_1^n i^{-r} &= \sum_1^{10} i^{-r} + 2\int_{11}^{n+1} x^{-r}\,dx - \int_{11}^{n+1}\int_x^{x+1} x^{-r}\,dx \\ &+ \frac{3}{4}\left[11^{-r} - \frac{(12)^{1-r}}{1-r} + \frac{(11)^{1-r}}{1-r} - n^{-r} + \frac{(n+1)^{1-r}}{1-r} - \frac{n^{1-r}}{1-r}\right] + e_2, \end{aligned}$$

* In the case $r = 1$, one must use the limiting form approached by the form in brackets as n approaches unity.

where

$$|e_2| \leqq \frac{1}{4}\left[.00669 - n^{-r} + \frac{(n+1)^{1-r}}{1-r} - \frac{n^{1-r}}{1-r}\right]$$

and

$$|e_2'| \leqq \frac{1}{4}(.00669) = .00167.$$

The limits of the errors in (19) are in general much smaller than the maximum, .00167. This is especially true when the series converges, . e., when $r>1$. We tabulate below a few such limits of errors.

r	.25	.333—	.4	.5	.75	1	2	3
limit of $\lvert e_2' \rvert$	.00150	.00163	.00167	.00164	.00134	.0009	.00017	.00002

ON GROUPS OF ORDER p^m, p BEING AN ODD PRIME NUMBER, WHICH CONTAIN AN ABELIAN SUBGROUP OF ORDER p^{m-1}.*

By H. A. Bender.

In the American Journal of Mathematics, vol. 24, pp. 395–8 is given a method of constructing all the groups of order p^m on the condition that those of order p^{m-1} are known, It is however pointed out in this article that not all the groups of order p^m obtained, in the manner given, from a particular group of order p^{m-1} are proved to be distinct, and in general, additional limitations must be imposed upon the generating elements so that all the groups thus generated will be distinct. The purpose of this article is to discuss some of these limitations.

We shall suppose G to be a non-abelian group of order p^m, and G_{m-1} to be the given subgroup of order p^{m-1}. Since every non-cyclic abelian group of order p^m is the direct product of independent cyclic groups we shall suppose a set of independent generators of G_{m-1} to be such that the group generated by any number of these generators has only the identity in common with the group generated by the remainig generators of the set. We shall suppose such a set of independent generators to be of orders $p^{\alpha_1}, p^{\alpha_2}, \cdots, p^{\alpha_n}$ $(\alpha_1 > \alpha_2 > \cdots > \alpha_n > 0)$, and that the number of the independent generators of these orders is $\lambda_1, \lambda_2, \cdots, \lambda_n$ respectively,

Suppose S_m to be an operator such that the group G is generated by G_{m-1} and S_m. Since the pth power of every operator in the major co-set $G_{m-1}S_m$ must be in G_{m-1}, we see from the well known formula

$$S_m^{-p} S_\alpha S_m^p = S_{\alpha-p} S_{\alpha-p+1}^p \cdots S_{\alpha-1}^p S_\alpha, \tag{1}$$

where $S_m^{-1} S_\alpha S_m = S_{\beta-1} S_\beta$ for $(\beta = \alpha, \alpha-1, \cdots, \alpha-p+1)$, that it is necessary for

$$S_{\alpha-p} S_{\alpha-p+1}^p \cdots S_{\alpha-1}^p = 1. \tag{2}$$

It follows from (2) that the pth commutator subgroup can not contain an independent generator. Since S_m must transform the operators of G_{m-1} in the same manner as they are transformed by isomorphisms of order p in the group of isomorphisms of G_{m-1}, then if in (1) we suppose

* Received October 6, 1926, in revised form Juli 1, 1927. Presented to the Society April 10, 1925, and September 9, 1926.

$S_{\alpha-\beta}$ ($\beta < p-1$) to be in the group generated by $S_{\alpha-\beta+1}, S_{\alpha-\beta+2}, \cdots S_{\alpha-1}, S_\alpha$, it must be in the group generated by $S_{\alpha-\beta+1}, \cdots, S^p_{\alpha-1}, S^p_\alpha$. Hence $S_{\alpha-p}\cdots, S^{(p)}_{\alpha-\beta}$ is an operator in the group generated by $S^p_{\alpha-\beta+1}, \cdots, S^{p^2}_{\alpha-1}, S^{p^2}_\beta$ and thus (2) can not be satisfied if the order of $S_{\alpha-1}$ is greater than p.

Hence there are two cases to be considered. In the first case we may suppose the pth commutator subgroup to be the identity, and in the second case we may suppose the pth commutator subgroup to be different from the identity. *All the operators of the commutator subgroups of the groups in the first case are of order p, except the identity, and the subgroup which is generated by the pth powers of the independent generators is contained in the central of every group of this class.* In case the pth commutator subgroup is different from the identity, it is evident from (2) that *it is necessary for G_{m-1} to contain at least $p-1$ independent generators either of the same order or the ratio of their orders is p.*

The four most important steps in the determination of the groups of order p^m which contain a given subgroup G_{m-1} are: first, the number of ways the central may be selected; second, for a given central H the number of ways the commutator subgroup may be selected; third, for a given central H and a given commutator subgroup K the possible ways of transforming G_{m-1} by the operators of the major co-set $G_{m-1}S_m$; fourth, the total number of groups which transform G_{m-1} in a given manner.

Consider first those groups in which the pth commutator subgroup is the identity. Since two groups are distinct if their centrals are not isomorphic, it follows that two groups are distinct if their centrals do not contain the same number of independent generators of like orders. Hence if the central omits λ independent generators of G_{m-1} ($\lambda = 1, 2, \cdots, \lambda_1+\lambda_2+\cdots+\lambda_n$), and if we suppose it to omit λ'_1 generators of order p^{α_1}, λ'_2 of order $p^{\alpha_2}, \cdots, \lambda'_n$ of order p^{α_n}, then all the groups containing a central of order $p^{m-1-\lambda}$ are distinct for all possible choices of λ'_i so that $\lambda = \sum_1^n \lambda'_i$ ($0 \leqq \lambda'_i \leqq \lambda_i$) ($i = 1, 2, \cdots, n$). If $\alpha_1 = 1$, then ($\lambda = 1, 2, \cdots, \lambda_1 - 1$).

We shall now study the manner in which the commutator subgroup may be chosen when a group has a given central H of order $p^{m-1-\lambda}$. It should be noted that two groups with a given central are distinct if their commutator subgroups do not include the same number of operators which are generated by the invariant independent generators, for in this case the commutator subgroups could not be made simply isomorphic. The λ generators of the commutator subgroup may be selected from $\lambda_i - \lambda'_i$ of the operators of order p which are generated by the $\lambda_i - \lambda'_i$ invariant

independent generators of order p^{α_i}, and from λ_i' of the operators of order p which are generated by the λ_i' non-invariant independent generators of order p^{α_i} $(i = 1, 2, \cdots, n)$. Hence if we suppose the λ generators of the commutator subgroup to contain γ_i operators from $\lambda_i - \lambda_i'$, and to contain γ_i' operators from the λ_i' operators $(i = 1, 2, \cdots, n)$, then all the groups having a given central H of order $p^{m-1-\lambda}$ $\left(\lambda = \sum_1^n \lambda_i'\right)$ are distinct for all possible choices of γ_i and γ_i' $(0 \leqq \gamma_i \leqq \lambda_i - \lambda_i',\ 0 \leqq \gamma_i' \leqq \lambda_i')$ so that $\lambda = \sum_1^n (\gamma_i + \gamma_i')$. If $\alpha_n = 1$, and if either $\lambda_j' = 0$ $(j < n)$ or $n = 1$, then λ_n' is to be replaced by $\lambda_n' - 1$.

We now come to the third step, and we shall discuss when two groups having a given central and a given commutator subgroup are distinct. Since two groups are distinct if the commutators formed by S_m and the independent generators of like orders are not simply isomorphic, it follows that all the groups containing a given central H and a given commutator subgroup K are distinct for all possible choices of γ_{ij} and γ_{ij}' $(i = 1, 2, \cdots, n,$ $j = 1, 2, \cdots, n)$ so that $\lambda = \sum_{i=1}^{n} \sum_{j=1}^{n} (\gamma_{ij} + \gamma_{ij}')$ $\left(\sum_{i=1}^{n} \gamma_{ij} = \gamma_j,\ \sum_{i=1}^{n} \gamma_{ij}' = \gamma_j'\right)$, where it is supposed that the commutators formed by S_m and the λ_i' non-invariant independent generators of order p^{α_i} contain γ_{i1} operators from γ_1, γ_{i1}' from γ_1', $\cdots$, γ_{in} from γ_n, γ_{in}' from γ_n' $\left[\lambda_i' = \sum_{j=1}^{n} (\gamma_{ij} + \gamma_{ij}')\right]$.

Thus far we have considered two groups to be distinct if one or more of the following conditions do not exist. Their centrals, or their commutator subgroups, or the subgroups which are generated by the commutators formed by S_m and the independent generators of like orders are simply isomorphic.

We have yet to discuss the distinct groups arising from a given central H and a given commutator subgroup K, and in which the subgroups which are generated by the commutators formed by S_m and the independent generators of like orders are isomorphic. The theorem concerning simply isomorphic groups must necessarily play an important role in the complete solution of this problem.

Suppose we wish to determine all the distinct groups of order p^5 which contain the abelian subgroup G_4 of order p^4 and type $(2, 2)$, and have a commutator subgroup of order p^2. Suppose two groups G_1, G_2 of order p^5 to be generated by G_4 and t_1, t_2 respectively. Suppose the generators of G_4 to be $S_1^{p^2} = S_2^{p^2} = 1$, and t_1, t_2 to be such that*

$$t_1^{-1} S_1 t_1 = S_1^{p\alpha_1} S_2^{p\alpha_2} S_1, \qquad t_2^{-1} S_1 t_2 = S_1^{px_1} S_2^{px_2} S_1,$$
$$t_1^{-1} S_2 t_1 = S_1^{p\beta_1} S_2^{p\beta_2} S_2, \qquad t_2^{-1} S_2 t_2 = S_1^{py_1} S_2^{py_2} S_2.$$

* G. Bagnera, Ann. di Mat., (3), vol. 1 (1898), p. 137.

If we wish to determine whether G_1 and G_2 are simply isomorphic by considering the most general isomorphism $\begin{pmatrix} t_1, & S_1, & S_2, \\ t_2^\mu, & S_1^{a_1} S_2^{a_2}, & S_1^{b_1} S_2^{b_2} \end{pmatrix}$, $(a_1 b_2 - a_2 b_1 \neq 0)$, then it will be necessary for the following congruences to be satisfied

$$\begin{aligned} a_1 \alpha_1 + b_1 \alpha_2 &\equiv \mu (a_1 x_1 + a_2 y_1) \bmod p, \\ a_1 \beta_1 + b_1 \beta_2 &\equiv \mu (b_1 x_1 + b_2 y_1) \bmod p, \\ a_2 \alpha_1 + b_2 \alpha_2 &\equiv \mu (a_1 x_2 + a_2 y_2) \bmod p, \\ a_2 \beta_1 + b_2 \beta_2 &\equiv \mu (b_1 x_2 + b_2 y_2) \bmod p. \end{aligned}$$

Hence to give a complete solution of this problem it is necessary to solve four congruences in four variables and five parameters, since we may suppose the α's and β's to have assigned values.

It is evident from this simple illustration that in general if we wish to determine a complete set of distinct groups it will be necessary to determine some restrictions to be placed on the simple isomorphisms.

THEOREM I. *If a group G contains an abelian subgroup of order p^{m-1}, generated by the μ independent generators $S_1, S_2, \cdots, S_\mu$ of order $p^{\alpha_1} (\alpha_1 > 1)$ $(\mu < p-1)$ which are such that a subgroup generated by less than these μ generators is not invariant under G, then all the operators in the major co-set $G_{m-2} S_1$ are conjugate under the group of isomorphisms of G, except those operators which might be transformed into some power of themselves.*

To prove this we may suppose $S_m^{-1} S_i S_m = s_{i+1} S_i (i = 1, 2, \cdots, \mu - 1)$ and $S_m^{-1} S_\mu S_m = s_1^{x_1} s_2^{x_2} \cdots s_\mu^{x_\mu} S_\mu$ where $S_i^{p^{\alpha_1 - 1}} = s_i$. We shall prove that the following isomorphism is possible. Let S_m correspond to itself, S_1 to $S_1^{a_1} S_2^{a_2} \cdots S_\mu^{a_\mu}$, and S_j to $S_1^{a_{1j}} S_2^{a_{2j}} \cdots S_\mu^{a_{\mu j}}$ $(j = 2, 3, \cdots, \mu)$, where it is necessary for $a_{ij} = a_{i-j+1} + A_\mu x_{i-j+2} + A_{\mu-1} x_{i-j+3} + \cdots + A_{\mu-j+2} x_i$ $(A_i = a_i + A_\mu x_{i+1} + A_{\mu-2} x_{i+2} + \cdots + A_{i+1} x_\mu)$ $(a_n = 0,\ x_n = 0,\ n \leqq 0)$ $(i = 1, 2, \cdots, \mu)$.

Such an isomorphism is possible if

$$a_i x_1 + \sum_{j=2}^{\mu} x_j a_{ij} = \sum_{k=1}^{i} A_{i-k+1} x_k. \tag{3}$$

It is easy to verify that this is a true relation by expanding and factoring x_k out of the left hand member.

In case there exists an n such that

$$n^\mu x_1 + n^{\mu-1} x_2 + \cdots + n x_\mu - 1 \equiv 0 \mod p, \tag{4}$$

then G contains a cyclic invariant subgroup generated by $S_1^{n x_1} S_2^{n^2 x_1 + n x_2} \cdots S_\mu$.

Thus if two groups in which $S_m^{-1}S_iS_m = s_{i+1}S_i$ $(i = 1, 2, \cdots, \mu-1)$, and in the one group $S_m^{-1}S_\mu S_m = s_1^{x_1}s_2^{x_2}\cdots s_\mu^{x_\mu}S_\mu$ and in the other group $S_m^{-1}S_\mu S_m = s_1^{y_1}s_2^{y_2}\cdots s_\mu^{y_\mu}S_\mu$ are isomorphic they can be made isomorphic by making S_1 isomorphic to itself, Furthermore, since these groups contain but one abelian subgroup of order p^{m-1} if the commutator subgroup is of order greater than p, then we may in addition make S_m isomorphic with some power of itself. *It should be noted, however, that these restrictions on isomorphisms do not necessarily hold in case* $\mu \geqq p-1$.

Similarly it can be shown that the above discussion holds if the μ independent generators are not all of the same order.

THEOREM II. *If a group G contains an abelian subgroup of order* p^{m-1}, *generated by the* μ *independent generators* $S_1, S_2, \cdots, S_\mu$ *of order* p^{α_1} $(\alpha_1 > 1)$ $(\mu < p-1)$ *which are such that they generate not more than one cyclic subgroup of order* p^{α_1} *the operators of which are transformed into the same pth power of themselves, and no subgroup generated by less than the* μ *generators is invariant under G, then there are* $p^{\mu-1}-1+d$ *distinct groups of order* p^m *which contain this abelian subgroup, d being the greatest common divisor of* μ *and* $p-1$.

It follows from the discussion above that we may consider the isomorphism $\begin{pmatrix} S_m, & S_1, & S_2, & S_3, & \cdots, & S_\mu \\ S_m^\alpha, & S_1, & S_2^\alpha, & S_3^{\alpha^2}, & \cdots, & S_\mu^{\alpha^{\mu-1}} \end{pmatrix}$. For the two groups to be isomorphic it is necessary for

$$x_i\alpha^{\mu-i+1} \equiv y_i \mod p \qquad (i = 1, 2, \cdots, \mu)\ (x_2 \not\equiv 0,\ y_1 \not\equiv 0).$$

Hence two of these groups will be distinct if the commutator of S_m *and* S_μ *are not generated by the same number of independent generators.* As α varies through the $p-1$ values we see that there are d distinct groups, d being the greatest common divisor of μ and $p-1$, when this commutator is generated by S_1, there are $(\mu-1)(p-1)$ distinct groups if this commutator is generated by S_1 and a second independent generator but not by S_1 alone, etc. If we would exclude the case $n^\mu x_1 + n^{\mu-1}x_2 + \cdots + nx_\mu - 1 = 0$, then G does not contain an invariant cyclic subgroup of order p^{α_1}. Theorem II does not include the cases in which the operators of two or more cyclic subgroups are transformed into the same power of themselves. In case the μ generators are each of a different order this gives the total number of distinct groups.

THEOREM III. *The total number of distinct groups of order* p^m *which contain a given abelian subgroup of order* p^{m-1} *generated by* μ *independent generators each of a different order greater than p* $(\mu < p-1)$ *is* $p^{\mu-1}-1+d$, *d being the greatest common divisor of* μ *and* $p-1$.

The fourth step consists in determining the number of distinct groups which transform G_{m-1} in the same manner, that is, to determine the possible powers of S_m. The pth power of any operator $S_m S_\alpha$ of the major co-set $G_{m-1} S_m$ is

$$(5) \qquad (S_m S_\alpha)^p = S_m^p S_{\alpha-p+1} S_{\alpha-p+2}^p \cdots S_\alpha^p$$

It follows from this that in the groups in which the $(p-1)$th commutator subgroup is the identity, the pth power of any operator of the major co-set $G_{m-1} S_m$ is

$$(S_m S_\alpha)^p = S_m^p S_\alpha^p,$$

and hence *any group in which the $(p-1)$th commutator subgroup is the identity which is generated by G_{m-1} and S_m is conformal with the abelian group of order p^m which is generated by G_{m-1} and an operator of the same order as S_m.* Furthermore, if in any of these groups the major co-set $G_{m-1} S_m$ contains an operator whose pth power is in the subgroup generated by the pth powers of the independent generators of G_{m-1}, then *this major co-set must contain an operator of order p. Hence we may suppose S_m to be an operator of order p, or to be such that its pth power is an invariant independent generator. If S_m is of order greater than p, then there may be two groups depending on whether some pth power of S_m is in the commutator subgroup or not.* All these groups will be distinct for different powers of S_m except in case G has a central of order p^{m-2}.

The above discussion is immediately applicable to the determination of the groups of order p^4 and p^5,* and those groups which contain operators of order p^{m-2} or p^{m-3}. For example, if G_{m-1} is of type (2,2) the congruences in the previous illustration reduce to

$$\mu^2 y_1 = \beta_1 \bmod p, \qquad \mu y_2 = \beta_2 \bmod p.$$

This gives all the groups of order p^5 with commutator subgroup of order p^2 $(p > 3)$ which contain the abelian subgroup of type (2,2), except the group in which every operator of this abelian subgroup is transformed into itself multiplied by its pth power.

We shall now study those groups in which the pth commutator subgroup is different from the identity. From (2) it is necessary for $S_{\alpha-p} S_{\alpha-p+1}^p \cdots S_{\alpha-1}^p = 1$. It will also be possible for the pth power of S_m to be commutative with S_α if

$$(2') \qquad S_{\alpha-p+1} S_{\alpha-p+2}^p \cdots S_\alpha^p = 1,$$

* Annals of Math., (2), vol. 28 (1927), pp. 61–71.

which results from (2) by replacing $\alpha-1$ by α since S_m^p is commutative with $S_{\alpha-1}$.

Let us suppose G_{m-1} to contain independent generators which are such that

$$S_m^{-1} S_\alpha S_m = S_{\alpha-1}^{p^a} S_\alpha, \quad S_m^{-1} S_\beta S_m = S_{\beta-1} S_\beta, \quad (\beta = \alpha-p+1, \cdots, \alpha-1),$$

then

$$S_m^{-p} S_\alpha S_m^p = (S_{\alpha-p} S_{\alpha-p+1}^p \cdots S_{\alpha-1}^p)^{p^a} S_\alpha.$$

It follows from this that if $S_{\alpha-1}$ is an independent generator of order p^{α_i}, then the commutator subgroup K must include the $p-2$ independent generators S_β $(\beta = \alpha-p+1, \cdots, \alpha-2)$ and $S_{\alpha-1}$ or some pth power of $S_{\alpha-1}$. Hence if $\alpha_i = \alpha_{i+1}+1$ and $\lambda_i+\lambda_{i+1} \geqq p-1$ then K may be chosen so as to contain $S_{\alpha-1}^{p^a}$ $(a = 0, 1, \cdots, \alpha_i-1)$ and $p-2-\lambda'_{i+1}$ $(\lambda_i \geqq p-1)$ independent generators of order p^{α_i}, or $\lambda_i-1-\lambda'_{i+1}$ in case $\lambda_i < p-1$, and the remaining λ'_{i+1} commutators to be of order $p^{\alpha_{i+1}}$ and, for each choice of K, S_α may be an independent generator of order equal to or greater than p^{α_i}. In case (2′) is satisfied we have the added restrictions that S_α must be of order p^{α_i} and the commutator formed by S_m and S_α can not be some pth power of $S_{\alpha-1}$. In either case these operators are transformed in but one way for a given K.

The pth power of the operators of the major co-set $G_{m-1} S_m$ is given by (5). It follows from this that the pth power of the operators of the co-set formed by multiplying S_m by the operators of the group generated by the $p-1$ independent generators, which are either of the same order or the ratio of their orders is p, is S_m^p. Hence *those groups in which the pth commutator subgroup is different from the identity can not be conformal with an abelian group if the commutator subgroup contains an operator of larger order than* S_m.

These groups are all distinct for different powers of S_m, and we may suppose S_m to be an operator of order p, or to be such that its pth power is an invariant independent generator, or suppose S_m to be of order p^2 and such that S_m^p be an invariant operator in the commutator subgroup.

OSCULATING DERIVATIVE OF A RULED SURFACE.*

By Dan Sun.

1. Introduction. Wilczynski† has shown that the theory of ruled surfaces may be based on the consideration of the system of two linear homogeneous differential equations of second order,

$$\text{(A)} \qquad \begin{aligned} x'' + p_{11}\,y' + p_{12}\,z' + q_{11}\,y + q_{12}\,z &= 0, \\ z'' + p_{21}\,y' + p_{22}\,z' + q_{21}\,y + q_{22}\,z &= 0, \end{aligned}$$

where

$$y' = dy/dx, \qquad y'' = d^2y/dx^2, \text{ etc.,}$$

and p_{ik}, q_{ik} $(i, k = 1, 2)$ are functions of an independent variable x. We shall speak of this system of differential equations as system (A) given by Wilczynski.

The four pairs of solutions (y_i, z_i), $(i = 1, 2, 3, 4)$, of the system (A) for which

$$\text{(1)} \qquad D \equiv \begin{vmatrix} y_1 & z_1 & y_1' & z_1' \\ y_2 & z_2 & y_2' & z_2' \\ y_3 & z_3 & y_3' & z_3' \\ y_4 & z_4 & y_4' & z_4' \end{vmatrix} \neq 0$$

constitute a fundamental set of simultaneous solutions.

These four pairs of solutions may be grouped into two quadruples (y_1, y_2, y_3, y_4) and (z_1, z_2, z_3, z_4), and we may interpret these two quadruples as the homogeneous coördinates of two points P_y and P_z in space. The coördinates of P_y and P_z being functions of x, as x varies P_y and P_z trace two curves C_y and C_z, called the integral curves of the system (A). Those points of these curves having the same values of x are said to be the corresponding points.

If the corresponding points P_y and P_z be joined by straight line L_{yz} the locus of L_{yz} will be a ruled surface S, called the integrating ruled surface of the system (A). If the determinant D is not identically equal to zero, this ruled surface is then a non-developable ruled surface.

Conversely, for every non-devolopable ruled surface there exists a system of differential equations of the form (A), and their coefficients can be deter-

* Received April 2, 1927; in revised form July 20, 1927.

† Wilczynski, Projective Differential Geometry of Curves and Ruled Surfaces, p. 126 et seq. We shall hereafter refer to this book as W.

mined by the parametric equations of two distinct curves on the ruled surface neither of which coincides with a generator of the ruled surface. These two curves may then be identified with the curves C_y and C_z in the preceding discussion, and may be called the directrix curves of the ruled surface.

At the corresponding points P_y and P_z of C_y and C_z, two corresponding osculating planes O_y and O_z are determined. These two planes will determine a line of intersection $\overline{L}_{yz}$. As x varies $\overline{L}_{yz}$ generates a ruled surface $\overline{S}$; I shall speak of $\overline{S}$ as the first osculating derivative of S with respect to directrix curves and to x, or simply the osculating derivative of S. In this paper I shall find the system of differential equations on which $\overline{S}$ is based and then discuss the properties of $\overline{S}$ in connection with the given surface S when the system (A) has constant coefficients.

2. **Generator of osculating derivative.** Any point $P_{\overline{y}}$ on the osculating plane O_y at P_y to C_y is given by

$$\overline{y} = y'' + \lambda_1 y' + \mu_1 y,$$

and any point $P_{\overline{z}}$ on the osculating plane O_z at P_z to C_z is given by

$$\overline{z} = z'' + \lambda_2 z' + \mu_2 z.$$

By means of (A) we have

$$\begin{aligned} \overline{y} &= (\lambda_1 - p_{11})\, y' - p_{12}\, z' + (\mu_1 - q_{11})\, y - q_{12}\, z, \\ \overline{z} &= -p_{21}\, y' + (\lambda_2 - p_{22})\, z' - q_{21}\, y + (\mu_2 - q_{22})\, z. \end{aligned}$$

For any point on the line of intersection of the osculating planes O_y and O_z,

$$\frac{\lambda_1 - p_{11}}{p_{21}} = \frac{p_{12}}{\lambda_2 - p_{22}} = \frac{\mu_1 - q_{11}}{q_{21}} = \frac{q_{12}}{\mu_2 - q_{22}}.$$

Let the common ratio be $\lambda(x)$; we have

$$\lambda_1 - p_{11} = \lambda p_{21}, \quad \lambda_2 - p_{22} = p_{12}/\lambda, \quad \mu_1 - q_{11} = \lambda q_{21}, \quad \mu_2 - q_{22} = q_{12}/\lambda.$$

Therefore

$$\begin{aligned} \overline{y} &= \lambda(p_{21}\, y' + q_{21}\, y) - (p_{12}\, z' + q_{12}\, z), \\ \overline{z} &= -(p_{21}\, y' + q_{21}\, y) + \frac{1}{\lambda}(p_{12}\, z' + q_{12}\, z). \end{aligned}$$

These show that the line of intersection of those two osculating planes is the line joining the points P_η and P_ζ, where η and ζ are given by

$$\eta = p_{21}\, y' + q_{21}\, y, \qquad \zeta = p_{12}\, z' + q_{12}\, z. \tag{2}$$

We can show that η and ζ are semi-covariants under the transformation of the independent variable x in the following way. We have seen that if we put $\xi = \xi(x)$, then*

$$\pi_{11} = \frac{1}{\xi'}\left(p_{11} + \frac{\xi''}{\xi'}\right), \qquad \pi_{12} = \frac{1}{\xi'} p_{12}, \qquad \pi_{21} = \frac{1}{\xi'} p_{21},$$
$$\pi_{22} = \frac{1}{\xi'}\left(p_{22} + \frac{\xi''}{\xi'}\right), \qquad \varkappa_{ik} = \frac{1}{\xi'^2} q_{ik}, \qquad (i, k = 1, 2),$$

where π_{ik}, $\varkappa_{ik}$ are the coefficients of the differential equations after the transformation of the independent variable. Also,

$$\frac{dy}{d\xi} = \frac{y'}{\xi'}, \qquad \frac{dz}{d\xi} = \frac{z'}{\xi'}.$$

Thus, if η and ζ are converted into $\bar{\eta}$ and $\bar{\zeta}$ by the transformation, we find

$$\bar{\eta} = \frac{1}{\xi'^2}\eta, \qquad \bar{\zeta} = \frac{1}{\xi'^2}\zeta,$$

which proves that η *and* ζ *are relative semi-covariants.* Therefore *the osculating derivative* $\bar{S}$ *is independent of the choice of the independent variable.*

Now we have, by (2), the osculating derivative $\bar{S}$ generated by the straight line $\bar{L}_{\eta\zeta}$ corresponding to the generator L_{yz} of S. Suppose the system (A) has been reduced to its semi-canonical form for which† $p_{ik} = 0$, $(i, k = 1, 2)$. Then C_y and C_z are two asymptotic curves‡ on the ruled surface S, and

$$\eta = q_{21} y, \qquad \zeta = q_{12} z.$$

Thus the generator $\bar{g}$ of $\bar{S}$ coincides with the corresponding generator g of S, and hence the ruled surface S and its osculating derivative $\bar{S}$ coincide. But the osculating plane of the asymptotic curve is the tangent plane of the ruled surface. Hence *the tangent plane at any point of a ruled surface contains the generator through this point.*

Differentiating (2) with respect to x and making the use of (A) we have

$$\begin{aligned} \eta' &= \alpha_{11} y' - p_{12} p_{21} z' + \alpha_{12} y - p_{21} q_{12} z, \\ \zeta' &= -p_{12} p_{21} y' + \alpha_{22} z' - p_{12} q_{21} y + \alpha_{21} z, \end{aligned} \tag{3}$$

where

$$\begin{aligned} \alpha_{11} &= p'_{21} - p_{11} p_{21} + q_{21}, & \alpha_{12} &= q'_{21} - p_{21} q_{11}, \\ \alpha_{21} &= q'_{12} - p_{12} q_{22}, & \alpha_{22} &= p'_{12} - p_{22} p_{12} + q_{12}. \end{aligned} \tag{4}$$

* W., p. 103.
† W., p. 114.
‡ W., p. 142.

Thus $\bar{D}$, the determinant of the ruled surface $\bar{S}$, is given by

$$\begin{aligned}\bar{D} &= (\eta', \zeta', \eta, \zeta)\\ &= (\alpha_{11}y' + \alpha_{12}y,\quad \alpha_{22}z' + \alpha_{21}z,\quad p_{21}y' + q_{21}y,\quad p_{12}z' + q_{12}z)\\ &= \alpha_1\alpha_2 D,\end{aligned}$$

where

$$\alpha_1 = p_{21}\alpha_{12} - q_{21}\alpha_{11}, \qquad \alpha_2 = p_{12}\alpha_{21} - q_{12}\alpha_{22}. \tag{5}$$

This shows that $\bar{D}$ vanishes if D, the determinant for S, vanishes. Therefore *the osculating derivative $\bar{S}$ is a developable if S itself is a developable. Moreover $\bar{S}$ is also a developable if $\alpha_1\alpha_2 \equiv 0$, i.e., if $\alpha_1 \equiv 0$ or $\alpha_2 \equiv 0$,* a case which we shall discuss later.

3. **Differential equations of osculating derivative.** Let us now set up the system of differential equations for the osculating derivative $\bar{S}$ of S. It will be, of course, a system of the form (A) between η and ζ.

From (3) we find

$$\eta + p_{21}\zeta = \alpha_{11}y' + \alpha_{12}y, \quad \zeta' + p_{12}\eta = \alpha_{22}z' + \alpha_{21}z.$$

Eliminating y' and z' with (2),

$$p_{21}(\eta' + p_{21}\zeta) - \alpha_{11}\eta = \alpha_1 y, \quad p_{12}(\zeta' + p_{12}\eta) - \alpha_{22}\zeta = \alpha_2 z. \tag{6}$$

Differentiating with respect to x and then eliminating y', z', y, z with (2) and (6); we obtain the system of the differential equations for the ruled surface $\bar{S}$, the osculating derivative of S,

$$\begin{aligned}&\eta'' + P_{11}\eta' + P_{12}\zeta' + Q_{11}\eta + Q_{12}\zeta = 0,\\ &\zeta'' + P_{21}\eta' + P_{22}\zeta' + Q_{21}\eta + Q_{22}\zeta = 0,\end{aligned} \tag{$\bar{\text{A}}$}$$

where

$$\begin{aligned}&P_{11} = p_{11} - \frac{d}{dx}\log\alpha_1, \quad P_{12} = p_{21}, \quad P_{21} = p_{12},\\ &P_{22} = p_{22} - \frac{d}{dx}\log\alpha_2, \qquad Q_{11} = \frac{\alpha_{11}}{p_{21}}\frac{d}{dx}\log\left(\frac{\alpha_1}{\alpha_{11}}\right) - \frac{\alpha_{12}}{p_{21}},\\ &Q_{12} = 2p'_{21} + q_{21} - p_{21}\frac{d}{dx}\log\alpha_1, \quad Q_{21} = 2p'_{12} + q_{12} - p_{12}\frac{d}{dx}\log\alpha_2,\\ &Q_{22} = \frac{\alpha_{22}}{p_{12}}\frac{d}{dx}\log\left(\frac{\alpha_2}{\alpha_{22}}\right) - \frac{\alpha_{21}}{p_{12}}.\end{aligned} \tag{7}$$

If the systems (A) and ($\bar{\text{A}}$) are reciprocal to each other, that is, if p_{ik} and q_{ik} are formed from P_{ik} and Q_{ik} just as P_{ik} and Q_{ik} are formed from p_{ik} and q_{ik}, we have the identities:

$$p_{11} \equiv P_{11} - \frac{d}{dx}\log A_1, \quad p_{12} \equiv P_{21}, \quad p_{21} \equiv P_{12},$$

$$p_{22} \equiv P_{22} - \frac{d}{dx}\log A_2, \qquad q_{11} \equiv \frac{A_{11}}{P_{21}}\frac{d}{dx}\log\left(\frac{A_1}{A_{11}}\right) - \frac{A_{12}}{P_{21}},$$

(8)

$$q_{12} \equiv 2P'_{21} + Q_{21} - P_{21}\frac{d}{dx}\log A_1, \quad q_{21} \equiv 2P'_{12} + Q_{12} - P_{12}\frac{d}{dx}\log A_2,$$

$$q_{22} \equiv \frac{A_{22}}{P_{12}}\frac{d}{dx}\log\left(\frac{A_2}{A_{22}}\right) - \frac{A_{21}}{P_{12}},$$

where A_i, A_{ik} are the quantities formed from the system $(\bar{A})$ in the same way as are α_i, α_{ik} from (A). Then we shall find, by means of (7),

$$\frac{d}{dx}\log \alpha_1 A_1 \equiv 0, \quad \frac{d}{dx}\log \alpha_2 A_2 \equiv 0,$$

$$\frac{A_{11}}{p_{12}}\frac{d}{dx}\log\left(\frac{A_1}{A_{11}}\right) + \frac{\alpha_{11}}{p_{21}}\frac{d}{dx}\log\left(\frac{\alpha_1}{\alpha_{11}}\right)$$

$$-\frac{1}{p_{12}}\left[2p''_{12} + q'_{12} - \frac{d}{dx}\left(p_{12}\frac{d}{dx}\log \alpha_2\right)\right] - \frac{q'_{21}}{p_{21}} \equiv 0,$$

(9)

$$4p'_{12} - p_{12}\frac{d}{dx}\log \alpha_2 A_1 \equiv 0, \quad 4p'_{21} - p_{21}\frac{d}{dx}\log \alpha_1 A_2 \equiv 0,$$

$$\frac{A_{22}}{p_{21}}\frac{d}{dx}\log\left(\frac{A_2}{A_{22}}\right) + \frac{\alpha_{22}}{p_{12}}\frac{d}{dx}\log\left(\frac{\alpha_2}{\alpha_{22}}\right)$$

$$-\frac{1}{p_{21}}\left[2p''_{21} + q'_{21} - \frac{d}{dx}\left(p_{21}\frac{d}{dx}\log \alpha_1\right)\right] - \frac{q'_{12}}{p_{12}} \equiv 0.$$

Hence we have the

THEOREM. *The relation between the systems* (A) *and* $(\bar{A})$ *is a reciprocal one if and only if*

$$\frac{d}{dx}(\alpha_1 A_1) = 0, \quad \frac{d}{dx}(\alpha_2 A_2) = 0,$$

(10)

$$\frac{d}{dx}\left(\frac{p^4_{12}}{\alpha_2 A_1}\right) = 0, \quad \frac{d}{dx}\left(\frac{p^4_{21}}{\alpha_1 A_2}\right) = 0$$

and (9_5), (9_6) *are identically satisfied.*

In case the coefficients of the equations of the system (A) are constants, the conditions (9_5), (9_6) and (10) are identically satisfied; and we have

$$P_{ik} = p_{ki}, \quad Q_{ik} = q_{ki}, \qquad (i, k = 1, 2).$$

Let us call the systems (A) and $(\bar{A})$ in case of constant coefficients (C) and $(\bar{C})$ respectively, then the system $(\bar{C})$ is given by

$$(\bar{C}) \qquad \begin{aligned} \eta'' + p_{11}\eta' + p_{21}\zeta' + q_{11}\eta + q_{21}\zeta &= 0, \\ \zeta'' + p_{12}\eta' + p_{22}\zeta' + q_{12}\eta + q_{22}\zeta &= 0, \end{aligned}$$

where p_{ik}, q_{ik} are the constant coefficients of the system (C). Hence we have the

THEOREM. *If the ruled surface S is given by the system* (C) *with constant coefficients, then* (C) *and the system of differential equations for its first osculating derivative $\bar{S}$ are reciprocal to each other.*

Furthermore if we define the second osculating derivative $\bar{\bar{S}}$ of S to be the first osculating derivative of $\bar{S}$, then (9_3), (9_6) and (10) are the necessary and sufficient conditions for $\bar{\bar{S}}$ to coincide with S. In case of constant coefficients, we have the

THEOREM. *Every ruled surface defined by the system* (C) *with constant coefficients has a second osculating derivative $\bar{\bar{S}}$ projectively equivalent to itself.*

4. **Directrix curves of S with plane branches.** The necessary and sufficient conditions that the directrix curves C_y and C_z of S given by the system (A) be plane curves, that is their differential equations be of third order, are* respectively

$$(11) \qquad \begin{aligned} \Delta_1 &= p_{12}s_{12} - q_{12}r_{12} = 0, & \Delta_2 &= p_{21}s_{21} - q_{21}r_{21} = 0; \\ r_{12} &= p_{12}(p_{11}+p_{22}) - p'_{12} - q_{12}, & r_{21} &= p_{21}(p_{11}+p_{22}) - p'_{21} - q_{21}, \\ s_{12} &= p_{11}q_{12} + p_{12}q_{22} - q'_{12}, & s_{21} &= p_{21}q_{11} + p_{22}q_{21} - q'_{21}. \end{aligned}$$

From (4) and (5), we can easily show that

$$(12) \qquad \alpha_1 \equiv -\Delta_2, \qquad \alpha_2 \equiv -\Delta_1.$$

These identities show that the osculating derivative $\bar{S}$ is a developable not only when S is a developable but also if one of the directrix curves C_y and C_z is a plane curve. Since the case $D \equiv 0$ is to be excluded here, we have

THEOREM. *The osculating derivative $\bar{S}$ is a developable if and only if either C_y or C_z is a plane curve.*

In his Chicago thesis, N. B. MacLean obtained the same result in case the directrix curves of S are flecnode curves.†

5. **Relations connecting derivative and osculating, derivative of ruled surface given by system** (C). We have seen that the differential equations of the derivative S' of the ruled surface S given by (A) are‡

* W., p. 230.

† N. B. MacLean, On certain surfaces related covariantly to a given ruled surface, p. 4, Diss. University of Chicago, 1923.

‡ W., p. 152.

$$\text{(A')}\qquad \begin{aligned} &\bar{y}''+\bar{p}_{11}\bar{y}'+\bar{p}_{12}\bar{z}'+\bar{q}_{11}\bar{y}+\bar{q}_{12}\bar{z}=0,\\ &\bar{z}''+\bar{p}_{21}\bar{y}'+\bar{p}_{22}\bar{z}'+\bar{q}_{21}\bar{y}+\bar{q}_{22}\bar{z}=0, \end{aligned}$$

where

$$\text{(13)}\qquad \bar{p}_{ik}=p_{ik}+\lambda_{ik},\qquad \bar{q}_{ik}=q_{ik}+\frac{1}{2}\sum_{j=1}^{2}\lambda_{ij}p_{jk};\qquad (i,k=1,2);$$

$$\text{(14)}\qquad \begin{aligned} 2J\lambda_{11}&=u_{21}v_{12}-u_{22}v_{11}, & 2J\lambda_{12}&=-u_{11}v_{12}+u_{12}v_{11},\\ 2J\lambda_{21}&=u_{21}v_{22}-u_{22}v_{21}, & 2J\lambda_{22}&=-u_{11}v_{22}+u_{12}v_{21};\\ & & J&=u_{11}u_{22}-u_{12}u_{21}; \end{aligned}$$

$$\text{(15)}\qquad \begin{aligned} u_{11}&=2p'_{11}-4q_{11}+p_{11}^2+p_{12}p_{21}, & u_{12}&=2p'_{12}-4q_{12}+p_{12}(p_{11}+p_{22}),\\ u_{21}&=2p'_{21}-4q_{21}+p_{21}(p_{11}+p_{22}), & u_{22}&=2p'_{22}-4q_{22}+p_{22}^2+p_{12}p_{21}; \end{aligned}$$

$$\text{(16)}\qquad \begin{aligned} v_{11}&=2u'_{11}+p_{12}u_{21}-p_{21}u_{12},\\ v_{12}&=2u'_{12}+(p_{11}-p_{22})u_{12}-p_{12}(u_{11}-u_{22}),\\ v_{21}&=2u'_{21}-(p_{11}-p_{22})u_{21}+p_{21}(u_{11}-u_{22}),\\ v_{22}&=2u'_{22}-p_{12}u_{21}+p_{21}u_{12} \end{aligned}$$

and

$$\text{(17)}\qquad \begin{aligned} w_{11}&=2v'_{11}+p_{12}v_{21}-p_{21}v_{12},\\ w_{12}&=2v'_{12}+(p_{11}-p_{22})v_{12}-p_{12}(v_{11}-v_{22}),\\ w_{21}&=2v'_{21}-(p_{11}-p_{22})v_{21}+p_{21}(v_{11}-v_{22}),\\ w_{22}&=2v'_{22}-p_{12}v_{21}+p_{21}v_{12}. \end{aligned}$$

If the ruled surface S is given by (C) with constant coefficients, then the formulas (13)–(17) hold by putting

$$p'_{ik}=0,\quad q'_{ik}=0,\quad u'_{ik}=0,\quad v'_{ik}=0,\qquad (i,k=1,2),$$

and we have

$$\text{(18)}\qquad v_{11}+v_{22}=0,\quad (p_{11}-p_{22})(v_{11}-v_{22})+2p_{21}v_{12}+2p_{12}v_{21}=0.$$

Let us call the system (A') in this case (C'). Then the osculating derivative of the derivative S' which I shall denote by $\overline{(S')}$ has the system of differential equations,

$$\overline{((C'))}\qquad \begin{aligned} &\bar{\eta}''+\bar{p}_{11}\bar{\eta}'+\bar{p}_{21}\bar{\zeta}'+\bar{q}_{11}\bar{\eta}+\bar{q}_{21}\bar{\zeta}=0,\\ &\bar{\zeta}''+\bar{p}_{12}\bar{\eta}'+\bar{p}_{22}\bar{\zeta}'+\bar{q}_{12}\bar{\eta}+\bar{q}_{22}\bar{\zeta}=0. \end{aligned}$$

Moreover the derivative of the osculating derivative $\bar{S}$ which I shall denote by $(\bar{S})'$ has the system of differential equations of the form,

$$((\bar{C})')\qquad \begin{aligned} &\bar{Y}''+\bar{P}_{11}\bar{Y}'+\bar{P}_{12}\bar{Z}'+\bar{Q}_{11}\bar{Y}+\bar{Q}_{12}\bar{Z}=0,\\ &\bar{Z}''+\bar{P}_{21}\bar{Y}'+\bar{P}_{22}\bar{Z}'+\bar{Q}_{21}\bar{Y}+\bar{Q}_{22}\bar{Z}=0. \end{aligned}$$

From $(\overline{C})$, we find

$$\overline{u}_{ik} = u_{ki}, \quad \overline{v}_{ik} = -v_{ki}, \quad \overline{w}_{ik} = w_{ki}, \qquad (i, k = 1, 2). \tag{19}$$

Thus we have

$$\overline{J} = J,$$

$$\begin{aligned} 2\overline{J}\,\overline{\lambda}_{11} &= -u_{12}v_{21} + u_{22}v_{11}, & 2\overline{J}\,\overline{\lambda}_{12} &= u_{11}v_{21} - u_{21}v_{11}, \\ 2\overline{J}\,\overline{\lambda}_{21} &= -u_{12}v_{22} + u_{22}v_{12}, & 2\overline{J}\,\overline{\lambda}_{22} &= u_{11}v_{22} - u_{21}v_{12} \end{aligned} \tag{20}$$

and

$$\begin{aligned} \overline{\lambda}_{11} + \overline{\lambda}_{22} + \lambda_{11} + \lambda_{22} &= 0, \\ 2J[\overline{\lambda}_{11} - \overline{\lambda}_{22} - \lambda_{11} + \lambda_{22}] &= 2v_{11}(u_{11} + u_{22}), \\ 2J[(\overline{\lambda}_{12} + \overline{\lambda}_{21}) - (\lambda_{12} + \lambda_{21})] &= (u_{11} + u_{22})(v_{12} + v_{21}), \\ 2J[(\overline{\lambda}_{12} - \overline{\lambda}_{21}) + (\lambda_{12} - \lambda_{21})] &= (u_{11} + u_{22})(v_{21} - v_{12}). \end{aligned} \tag{20'}$$

Hence

$$\overline{P}_{ik} = p_{ki} + \overline{\lambda}_{ik}, \qquad \overline{Q}_{ik} = q_{ki} + \frac{1}{2}\sum_{j=1}^{2} \overline{\lambda}_{ij}\, p_{kj}, \qquad (i, k = 1, 2). \tag{21}$$

Now $(\overline{S})'$ is projectively equivalent to $(\overline{S'})$ if and only if

$$\overline{P}_{ik} = \overline{p}_{ki}, \qquad \overline{Q}_{ik} = \overline{q}_{ki}, \qquad (i, k = 1, 2).$$

Thus we have

$$\overline{\lambda}_{ik} = \lambda_{ki}, \qquad \sum_{j=1}^{2} \overline{\lambda}_{ij}\, p_{jk} = \sum_{j=1}^{2} \lambda_{ij}\, p_{jk}, \qquad (i, k = 1, 2). \tag{22}$$

From (20′) and (22), we get

$$\begin{aligned} &\lambda_{11} + \lambda_{22} = 0, \quad \lambda_{12}p_{21} - \lambda_{21}p_{12} = 0, \quad 2\lambda_{11}p_{12} - \lambda_{12}(p_{11} - p_{22}) = 0, \\ &v_{11}(u_{11} + u_{22}) = 0, \quad (u_{11} + u_{22})(v_{12} + v_{21}) = 0, \quad (u_{21} + u_{22})(v_{12} - v_{21}) = 0. \end{aligned} \tag{23}$$

Excluding the case $u_{11} + u_{22} = 0$ we have the necessary and sufficient conditions that $(\overline{S})'$ is projectively equivalent to $(\overline{S'})$:

$$v_{ik} = 0, \qquad (i, k = 1, 2). \tag{24}$$

Thus,

$$w_{ik} = 0, \qquad (i, k = 1, 2),$$

$$\theta_9 = \begin{vmatrix} u_{11} - u_{22} & u_{12} & u_{21} \\ v_{11} - v_{22} & v_{12} & v_{21} \\ w_{11} - w_{22} & w_{12} & w_{21} \end{vmatrix} = 0,$$

$$\theta_{10} = (v_{11}v_{22} - v_{12}v_{21})^2\theta_4 = 0$$

and also all of the minors of the second order in the determinant θ_9 vanish; while

$$\theta_4 = (u_{11} - u_{22})^2 - 4u_{12}u_{21} \neq 0$$

for general case. Hence we have the necessary condition:*

If $(\overline{S})$ is projectively equivalent to $(\overline{S})'$, then S must belong to a linear congruence; its directrices are in general distinct.

This condition is also sufficient. For, if S belongs to a linear congruence with distinct directrices, then all of the minors of the second order in the determinant θ_9 vanish and we have

$$(25)\quad \frac{v_{11}-v_{22}}{u_{11}-u_{22}} = \frac{v_{12}}{u_{12}} = \frac{v_{21}}{u_{21}} = k, \qquad \frac{w_{11}-w_{22}}{v_{11}-v_{22}} = \frac{w_{12}}{v_{12}} = \frac{w_{21}}{v_{21}} = \bar{k}.$$

From (16) and (25_1), for which $u'_{ik} = 0$, we have

$$(26)\quad \begin{aligned} &k(u_{11}-u_{22}) + 2p_{21}u_{12} - 2p_{12}u_{21} = 0,\\ &p_{12}(u_{11}-u_{22}) + [k-(p_{11}-p_{22})]u_{12} = 0,\\ &p_{21}(u_{11}-u_{22}) - [k+(p_{11}-p_{22})]u_{21} = 0. \end{aligned}$$

Since $u_{11}-u_{22}$, u_{12}, u_{21} are not all zero, otherwise $\theta_4 = 0$, the two directrices of the linear congruence coincide and S is a quadric. It follows that

$$\begin{vmatrix} k & 2p_{21} & -2p_{12} \\ p_{12} & k-(p_{11}-p_{22}) & 0 \\ p_{21} & 0 & -k-(p_{11}-p_{22}) \end{vmatrix} = 0,$$

$$(27)\qquad k[k^2-(p_{11}-p_{22})^2-4p_{12}p_{21}] = 0.$$

If $k = 0$, then $v_{11}-v_{22} = v_{12} = v_{21} = 0$; and from the first equation of (18), we have $v_{ik} = 0$, $(i, k = 1, 2)$. Hence the proof is complete.

If $k \neq 0$, then $v_{11}-v_{22}$, v_{12}, v_{21} are not all zero. From (25_2), we have

$$(26')\quad \begin{aligned} &\bar{k}(v_{11}-v_{22}) + 2p_{21}v_{12} - 2p_{12}v_{21} = 0,\\ &p(v_{11}-v_{22}) + [\bar{k}-(p_{11}-p_{22})]v_{12} = 0,\\ &p(v_{11}-v_{22}) - [\bar{k}+(p_{11}-p_{22})]v_{21} = 0. \end{aligned}$$

Hence

$$(27')\qquad \bar{k}[\bar{k}^2-(p_{11}-p_{22})^2-4p_{12}p_{21}] = 0.$$

Again, from the second equation of (18) and the last two equations of (26′), we have

* W., p. 168.

$$\begin{vmatrix} p_{11}-p_{22} & 2p_{21} & 2p_{12} \\ p_{12} & \bar{k}-(p_{11}-p_{22}) & 0 \\ p_{21} & 0 & -\bar{k}-(p_{11}-p_{22}) \end{vmatrix} = 0,$$

$$(27'') \qquad (p_{11}-p_{22})[\bar{k}^2-(p_{11}-p_{22})^2-4p_{12}p_{21}] = 0.$$

Since the equations (27′) and (27″) are satisfied simultaneously, it follows that $\bar{k} \neq 0$ and we have $\bar{k}^2-(p_{11}-p_{22})^2-4p_{12}p_{21}=0$; therefore $k=\bar{k}$. Continuing the process and supposing $t_{11}=2w'_{11}+p_{12}w_{21}-p_{21}w_{12}$, etc., then

$$\frac{t_{11}-t_{22}}{w_{11}-w_{22}} = \frac{t_{12}}{w_{12}} = \frac{t_{21}}{w_{21}} = k.$$

Hence k must be a seminvariant,* and so is k^2. But k^2 is not a seminvariant, for it can not be expressed as a function of u_{ik}.† This contradiction proves that $k=0$ and the condition is sufficient.

We may recapitulate the results in the following

THEOREM. *The necessary and sufficient condition for the osculating derivative* $(\overline{S'})$ *of the derivative* S' *of a ruled surface* S *given by the system* (C) *with constant coefficients, for which* $u_{11}+u_{22} \neq 0$, *to be projectively equivalent to the derivative* $(\overline{S})'$ *of the osculating derivative* $\overline{S}$ *is that* S *belongs to a linear congruence with distinct directrices.*

In case

$$(28) \qquad u_{11}-u_{22} = u_{12} = u_{21} = 0,$$

the equations (24) are identically satisfied, and we have‡ the

COR. *If* S *is a ruled surface of second order given by the system* (C) *with constant coefficients, then* $(\overline{S'})$ *is projectively equivalent to* $(\overline{S})'$.

6. **Fundamental invariants.** The fundamental invariants of the system (A) are given by§

$$\begin{aligned} \theta_4 &= I^2-4J, \\ \theta_{4\cdot 1} &= 8\theta_4\theta''_4-9\theta'^2_4+8I\theta_4^2, \\ (29) \qquad \theta_{10} &= (I^2-4J)(K^2-I'^2)+(II'-2J')^2, \\ \theta_9 &= \begin{vmatrix} u_{11}-u_{22} & u_{12} & u_{21} \\ v_{11}-v_{22} & v_{12} & v_{21} \\ w_{11}-w_{22} & w_{12} & w_{21} \end{vmatrix}, \end{aligned}$$

* W., p. 101.
† W., p. 96.
‡ W., p. 139.
§ W., Ch. IV.

where

$$(30)\quad I = u_{11}+u_{22},\quad J = u_{11}u_{22}-u_{12}u_{21},\quad K = v_{11}v_{22}-v_{12}v_{21},$$

and u_{ik}, v_{ik}, w_{ik} are given by (15), (16) and (17).

For the system (C) with constant coefficients, we have the fundamental invariants:

$$\theta_4 = I^2-4J,\quad \theta_{4\cdot 1} = 8I\theta_4^2,\quad \theta_{10} = K^2(I^2-4J),$$

$$(29')\qquad \theta_9 = \begin{vmatrix} u_{11}-u_{22} & u_{12} & u_{21} \\ v_{11}-v_{22} & v_{12} & v_{21} \\ w_{11}-w_{22} & w_{12} & w_{21} \end{vmatrix}$$

with

$$p'_{ik} = q'_{ik} = 0,\quad u'_{ik} = v'_{ik} = 0,\qquad (i, k = 1, 2),$$

in (15), (16) and (17).

Denote the fundamental invariants of the system $(\overline{C})$ by $\overline{\theta}_4$, $\overline{\theta}_{4\cdot 1}$, $\overline{\theta}_{10}$ and $\overline{\theta}_9$. Since

$$(31)\qquad \overline{I} = I,\quad \overline{J} = J,\quad \overline{K} = K.$$

Hence

$$(32)\qquad \overline{\theta}_4 = \theta_4,\quad \overline{\theta}_{4\cdot 1} = \theta_{4\,1},\quad \overline{\theta}_{10} = \theta_{10},\quad \overline{\theta}_9 = \theta_9.$$

These equalities show that the fundamental invariants of the two systems (C) and $(\overline{C})$ are same. Hence we have* the

THEOREM. *The ruled surface S given by the system* (C) *with constant coefficients and its osculating derivative $\overline{S}$ are projectively related.*

Hence the theorem concerning the projective equivalence of S and $\overline{S}$ on p. 100 is an immediate consequence of this theorem.

CHICAGO, ILL.

* W., p. 143.

CHAINS OF CONGRUENCES FOR THE NUMERATORS AND DENOMINATORS OF THE BERNOULLI NUMBERS.*

By Jessie L. Bell.

1. Even-suffix notation for Bernouilli's numbers is used: $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$, $B_8 = -1/30$, etc., and we write $B_{2n} = N_{2n}/D_{2n}$, N_{2n}, D_{2n} being the numerator and denominator of B_{2n}, so that N_{2n} and D_{2n} are relatively prime integers. (In the table B_{2n} is used to indicate the numerator, N_{2n}, only of B_{2n}.) If n is odd N_{2n} and D_{2n} are both taken positive; if n is even we take N_{2n} positive and D_{2n} negative. Write, with Lucas, G_{2n} for the $2n$th Gennochi number defined by

$$G_{2n} = 2(1-2^{2n})\,B_{2n}.$$

Then it was shown by Lucas (Théorie des Nombres, pp. 434–5) that G_{2n} is an integer, and that if p is an odd prime,

$$2n\,G_{2n+p-1} \equiv (2n-1)\,G_{2n} \mod p.$$

Substituting for G its value in terms of B and replacing B by N/D, we infer that (since $2^{p-1} \equiv 1 \bmod p$ by Fermat's theorem),

$$2n(4^n-1)\,N_{2n+p-1}\,D_{2n} \equiv (2n-1)(4^n-1)\,N_{2n}\,D_{2n+p-1} \mod p;$$

for p an odd prime. If 4^n-1 is not divisible by p, we may divide out 4^n-1 from the preceding congruence, and get

$$2n\,N_{2n+q-1}\,D_{2n} \equiv (2n-1)\,N_{2n}\,D_{2n+q-1} \mod q,$$

where now q is any odd prime which is not a divisor of 4^n-1. For given values of q, n it may happen that this congruence is further reducible, viz., both members may contain a common factor prime to q. Or, for n given, $2n\,D_{2n}$ and $(2n-1)\,N_{2n}$ may both be divisible by one of the prime divisors of 4^n-1. Taking account of all possible reductions we can state the following

Theorem. *For each positive integral value of n there exist constant positive integers a, b not both divisible by a prime factor of $2^n \pm 1$ such that*

$$a\,N_{q+2n-1} \equiv b\,D_{q+2n-1} \mod q$$

* Received July 5, 1927.

for every prime q except possibly a finite number of primes which are divisors of $2^n \pm 1$.

2. The numbers B_{2n} have been calculated by J. C. Adams up to B_{124}, and the prime factors of $2^n \pm 1$ for $n \leqq 63$ were collected by Lucas in the *American Journal of Mathematics*, vol. 1, pp. 239–240. It will be of interest to give the sets of congruences which can be inferred from these lists, for each such set provides a check upon all the Bernoulli numbers of a given class. Thus, putting $n = 8$, we get the congruences

$$a N_{q+15} \equiv b D_{q+15} \mod q,$$

holding for all primes q except possibly 3, 5, 17, 257, (17 tests out not excluded as shown by the table below) and (see below) $a = -544$, $b = +3617$. Hence the odd primes *not* excluded being 7, 11, 13, 17, 19, 23, 29, 31, ···, 251, 263, ··· to ∞, we can use this congruence to check B_{2n} for the following values of $2n$, 22, 26, 28, 32, 34, 38, 44, 46, ···, 266, 278, ··· to ∞.

And similarly each case $n = 1, 2, 3, \cdots$ furnishes its own series of related Bernoulli numbers, the series corresponding to n being the sequence of primes (with a certain possible finite number of omissions) each increased by $2n-1$.

To construct the reduced congruence for a given n, we first find the possible excluded prime moduli q by writing down the distinct prime factors of the Fermat numbers $2^n \pm 1$ from Lucas' tables. Let r be one of these primes. Then if it is a divisor of both $2n D_{2n}$ and $(2n-1) N_{2n}$, the congruence $2n D_{2n} N_{2n+r-1} \equiv (2n-1) N_{2n} D_{2n+r-1} \mod r$ is reducible in the sense stated in § 1, and we leave r in the excluded primes to be tested for possible rejection from the list by inspection. The last is done thus: if for the given n the value of $2n+r-1$ does not exceed 124 (the limit of Adams' table) we can test by actual trial the divisibility by r of the two sides of the congruence in the theorem.

3. Before giving the table we offer a detailed example of the method of its construction. Take $n = 11$. Then the possibly excluded prime moduli will be divisors of $2^{11} \pm 1$, and hence from Lucas' table, 3, 23, 89, 683. The congruence (§1)

$$2n D_{2n} N_{q+2n-1} \equiv (2n-1) N_{2n} D_{q+2n-1} \mod q$$

for $n = 11$ therefore becomes

$$22 D_{22} N_{q+21} \equiv 21 N_{22} D_{q+21} \mod q,$$

valid for all primes except *possibly* $q = 3, 23, 89, 683$, or say

$$22\,D_{22}N_{q+21} \equiv 21\,N_{22}D_{q+21} \mod q \neq 3,\ 23,\ 89,\ 683,$$

which is certainly true, but perhaps too restricted. From Adams' table we have, substituting in the values of D_{22} and N_{22} and factoring,

$$2\cdot 11\cdot 3\cdot 2\cdot 23\,N_{q+21} \equiv 3\cdot 7\cdot 11\cdot 77683\,D_{q+21} \mod q \neq 3,\ 23,\ 89,\ 683.$$

Now both sides of the congruence will divide out by 3 and 11 and we can thus reduce the congruence, *provided* we include 3 and 11 among the excluded primes. Hence finally

$$2^2\cdot 23\,N_{q+21} \equiv 7\cdot 77683\,D_{q+21} \mod q \neq 3,\ 11,\ 23,\ 89,\ 683.$$

In some cases we can go farther. Where $q+21 \lesseqgtr 124$ from Adams' table, we can substitute in the five different values of N_{q+21} and D_{q+21} satisfying this condition and ask whether the congruence holds for the respective q's. We find that the congruence holds for $q = 23$ and $q = 89$. The prime 683 is too large to test; so that there are left only three excluded primes for which the congruence possibly does not hold, 3, 11, 683. This congruence then gives us information regarding all the following Bernoulli numbers B_{2n} which we obtain by first writing down the primes excluding 3, 11 and the doubtful 683;

$$5,\ 7,\ 13,\ 17,\ 19,\ 23,\ \cdots,\ 2729,\ 2741,\ \cdots,\ 8179,\ 8209,\ \cdots$$

then adding 21 to each:

$$2n = 26,\ 28,\ 34,\ 38,\ 40,\ 44,\ \cdots,\ 2750,\ 2762,\ \cdots,\ 8200,\ 8230,\ \cdots.$$

That is, we have the infinite chain of congruences

$$92\,N_{26} \equiv 543781\,D_{26} \mod 5, \qquad 92\,N_{28} \equiv 543781\,D_{28} \mod 7,$$
$$92\,N_{34} \equiv 543781\,D_{34} \mod 13,\ \cdots,\ 92\,N_{8230} \equiv 543781\,D_{8230} \mod 8209.$$

4. The table suggests an interesting and difficult question: Is it possible to find a finite number of these chains which are such that they provide congruence tests for all but a finite number (which may be none) of the Bernoulli numbers? If this is not possible, what is the average number of gaps in a long stretch?

The writer is indebted to E. T. Bell for suggestions in this introduction. A Monroe machine has been used for the table, the work twice checked. The Bernoulli numbers in the table are all Bernoulli numerators. The numbers starred are those too large to be tested by Adams' table, given in *Crelle's Journal*, vol. 85, p. 270–2, (1878).

$$a N_{q+2n-1} \equiv b D_{q+2n-1} \mod q.$$

n	a	b	*Excluded Primes*
1	$2^2 \cdot 3$	1	
2	$(-1)\, 2^3 \cdot 5$	1	3
3	$2^2 \cdot 3^2 \cdot 7$	5	3^2
4	$(-1)\, 2^4 \cdot 3 \cdot 5$	7	17 divides both sides with a residue of 2
5	$2^2 \cdot 11$	3	3, 31; 5 divides both sides with a residue of 3
6	$(-1)\, 2^3 \cdot 5 \cdot 3^2 \cdot 7 \cdot 13$	$11 \cdot 691$	3^2
7	$2^2 \cdot 3$	13	7, 43, 127*
8	$(-1)\, 2^5 \cdot 17$	3617	3, 5, 257*
9	$2^2 \cdot 3^3 \cdot 7 \cdot 19$	$17 \cdot 43867$	3^3; 73 leaves a residue of 24 on both sides of the congruence
10	$(-1)\, 2^3 \cdot 3 \cdot 5^2 \cdot 11$	$19 \cdot 174611$	41, 31, 5^2
11	$2^2 \cdot 23$	$7 \cdot 77683$	3, 11, 683*; 89 leaves 44 on each side
12	$(-1)\, 2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 13$	$23 \cdot 236364091$	3^2, 241*; 17 leaves 5 on each side
13	$2^2 \cdot 3$	$5^2 \cdot 657931$	2731*, 8191*; 13 leaves 7 on each side
14	$(-1)\, 2^3 \cdot 5 \cdot 7 \cdot 29$	$3^2 \cdot 23749461029$	3, 43, 113*, 127*
15	$2^2 \cdot 3^2 \cdot 7 \cdot 11 \cdot 31$	$29 \cdot 1723168255201$	3^2, 151*, 331*
16	$(-1)\, 2^6 \cdot 3 \cdot 5$	31	257*, 65537*
17	2^2	$11 \cdot 151628697551$	3, 43691*, 131071*
18	$(-1) 2^3 \cdot 3^3 \cdot 13 \cdot 19 \cdot 37$	2631527155305347737 3	3^3, 5, 7, 73, 109*
19	$2^2 \cdot 3$	$37 \cdot 154210205991661$	19, 174763*, 524287*
20	$(-1)\, 2^4 \cdot 5^2 \cdot 11 \cdot 41$	$13 \cdot 26108271849644912 2051$	3, 5^2, 31, 61681*; 17 leaves a residue of 4 on each side
21	$2^2 \cdot 3^2 \cdot 7^2 \cdot 43$	$41 \cdot 1520097643918070802691$	3^2, 7^2, 127*, 337*, 5419*
22	$(-1)\, 2^3 \cdot 3 \cdot 5 \cdot 11 \cdot 23$	$43 \cdot 27833269579301024235023$	89*, 397*, 683*, 2113*
23	$2^2 \cdot 47$	$5 \cdot 3 \cdot 59645111159391216327796 1$	3, 178481*, 2796203*
24	$(-1)\, 2^5 \cdot 3^2 \cdot 5 \cdot 7 \cdot 13 \cdot 17$	$47 \cdot 56094033689978176862491275 47$	3^2, 97*, 241*, 257*, 673*
25	$2^2 \cdot 3 \cdot 5^2 \cdot 11$	$7^2 \cdot 49505720524107964821247752 5$	31, 251*, 601*, 1801*, 4051*
26	$(-1)\, 2^3 \cdot 5 \cdot 53$	$17 \cdot 80116571813548995734792499185 3$	3, 157*, 1613*, 2731*, 8191*
27	$2^2 \cdot 3^4 \cdot 7 \cdot 19$	$53 \cdot 29149963634884862421418123812 691$	3^4*, 73*, 87211*, 262657*

n	a	b	*Excluded Primes*
28	$(-1)\,2^4\cdot 3\cdot 29$	$11\cdot 247939292931322675368541573 9663229$	5, 17, 43, 113*, 127*, 15790321*
29	$2^2\cdot 29\cdot 59$	$19\cdot 844836133488800418620467759 94036021$	3, 233*, 1103*, 2089*, 3033169*
30	$(-1)\,2^3\cdot 3^2\cdot 5^2\cdot 11\cdot 13\cdot 7\cdot 31\cdot 61$	$59\cdot 12152331404837555720403049940 79820246041491$	3^2, 5^2, 151*, 331*, 1321*; 41 leaves a residue of 18 on each side
31	$2^2\cdot 3\cdot 31$	$61\cdot 12300585434086858541953039857 403386151$	715827883*, 2147483647*
32	$(-1)\,2^7\cdot 5\cdot 17$	$3\cdot 7\cdot 10678383014786652988638544 4979142647942017$	3, 257*, 641*, 65537*, 6700417*
33	$2^2\cdot 3^2\cdot 7\cdot 11\cdot 23\cdot 67$	$5\cdot 13\cdot 14726000221263356540516194285519323422418991 01$	3^2, 67*, 89*, 683*, 20857*, 599479*
34	$(-1)\,2^3\cdot 5\cdot 3$	$67\cdot 78773130858718728141909149208 474606244347001$	137*, 953*, 26317*, 43691*, 131071*.
35	$2^2\cdot 11\cdot 71$	$23\cdot 15053813473333670038030765673 778572085114381602 35$	3, 31, 43, 71*, 127*, 281*, 86171*, 122921
36	$(-1)\,2^4\cdot 3^3\cdot 5\cdot 7\cdot 13\cdot 19\cdot 37\cdot 73$	$71\cdot 58279549616699441104382772446 4106736528248830184426042 9$	3^3, 73*, 109*, 241*, 433*, 38737*; 17 leaves a residue of 5 on both sides
37	$2^2\cdot 3\cdot 37$	$73\cdot 34152417289221168014330073731 4726351866883077830 87$	223*, 1777*, 25781083*, 616318177*
38	$(-1)\,2^3$	$5\cdot 24655088825935372707687196040 5851999043652678288658 01$	3, 5, 229*, 457*, 174763*, 524287*, 525313*
39	$2^2\cdot 3^2\cdot 13\cdot 79$	$11\cdot 41484636557540082829517903554954207349219937537240048348 7$	3^2, 7, 79*, 2731*, 8191*, 121369*, 22366891*
40	$(-1)\,2^5\cdot 5^2\cdot 3\cdot 11\cdot 17\cdot 41$	$79\cdot 46037842994794576469355749690190468497942578727512889196 56867$	5^2, 31, 257*, 61681*, 4278255361*
41	$2^2\cdot 83$	$3^3\cdot 16770141491851458368231545097862699002077360275702534148 81613$	3, 83*, 13367*, 164511353*, 8831418697*
42	$(-1)\,2^3\cdot 3^2\cdot 5\cdot 7^2\cdot 13\cdot 29\cdot 43$	$13\cdot 20245761959352903602311311601117310099899173911980908772 81083932477$	3^2, 7^2, 43*, 113*, 127*, 337*, 1429*, 5419*, 14449*
43	$2^2\cdot 3$	$5\cdot 17\cdot 66071461941767865357384784742626149627783068665338 8931761996983$	431*, 9719*, 2099863*, 2932031007403*,
44	$(-1)\,2^4\cdot 5\cdot 11\cdot 23\cdot 89$	$29\cdot 13114264886740175079955114240193118433457502755720286442 96919890574047$	3, 89*, 353*, 397*, 683*, 2113*, 2931542417*; 17 divides both sides with a residue of 9

n	a	b	*Excluded Primes*
45	$3^3 \cdot 2^2 \cdot 7 \cdot 11 \cdot 19 \cdot 31$	89 · 1179057279021082799884123351 24921508377525494966964711623 1545215727922535	3^3, 73*, 151*, 331*, 631*, 23311*, 18837001*
46	$(-1)\, 2^3 \cdot 3 \cdot 5 \cdot 47$	91 · 1295585948207537527989427828 53857674965934148371943514302 3316326829946247	47*, 277*, 1013*, 1657*, 30269*, 178481*, 2796203*
47	2^2	31 · 1220813806579744469607301679 41320120395850841520269662143 621510528464944 7	3, 47*, 283*, 2351*, 4513*, 13264529*, 165768537521*
48	$(-1)\, 2^6 \cdot 3^2 \cdot 7 \cdot 13 \cdot 17 \cdot 97$	19 · 2116004495972665130975977281 09824233673043954389060234150 63873342005668349987259	3^3, 5, 97*, 193*, 241*, 257*, 673*, 65537*, 22253377*
49	$2^2 \cdot 3$	97 · 6790826067290549562405111754 64036056073421957285044875090 739612499929470582 39	7^2*, 43*, 127*, 4363953127297*, 4432676798593*
50	$(-1)\, 2^3 \cdot 5^3 \cdot 101$	3 · 9459803781912212529522743306 9493721872702841533066936133 385696204311395415197247711	3, 11, 5^3*, 31*, 41*, 101*, 251*, 601*, 1801*, 4051*, 8101*, 268501*
51	$2^3 \cdot 3^2 \cdot 7 \cdot 17 \cdot 103$	101 · 3204019410860907078243020782 11624177549181719715271745 0679002501086861530836678158 791	3^2, 103*, 307*, 2143*, 2857*, 6529*, 11119*, 43691*, 131071*
52	$(-1)\, 2^4 \cdot 3 \cdot 5 \cdot 53$	103 · 3195336313638300112871033527 96174274671189606078272738 32710347016284956836554972 1224053	53*, 157*, 1613*, 2731*, 8191*, 858001*, 308761441*; 17 divides both sides with a residue of 10
53	$2^2 \cdot 107$	5 · 7 · 3637390317261741440815182 01515934271692312986405816 90038930816378281879873386 202346572901	3, 107*, 6361*, 69431*, 20394401*, 28059810762433*
54	$(-1)\, 2^3 \cdot 3^4 \cdot 5 \cdot 7 \cdot 13 \cdot 19 \cdot 37 \cdot 109$	107 · 3469342247847828789552088659 32385254139976678576049114 68700058913715012663197248975 92306597338057	3^4*, 19*, 37*, 73*, 109*, 87211*, 246241*, 262657*, 279073*
55	$2^2 \cdot 3 \cdot 11^2 \cdot 23$	109 · 7645992940484742892248134246 72434750052875241341230790 66835938707597976062695857 79977930217515	23*, 31*, 89*, 11^2*, 683*, 881*, 2971*, 3191*, 201961*, 48912491*
56	$(-1)\, 2^5 \cdot 5 \cdot 7 \cdot 17 \cdot 29 \cdot 113$	37 · 2650879602155099713352597214 68516201444315149919250989645 17884276809667565148755153 6781203552600109	3, 17*, 29*, 43*, 113*, 127*, 257*, 5153*, 15790321*, 54410972897*

n	a	b	*Excluded Primes*
57	$2^2 \cdot 3^2 \cdot 7$	113 · 217378323193691633333107610 866529914757211566790908313 608061101149336054842345936 50904188618562649	3^2, 571*, 32377*, 174763*, 524287*, 1212847*, 160465489*
58	$(-1)\,2^3 \cdot 3 \cdot 59$	23 · 309553916571842976912513458033841416869004128064329844245504045721008957524571968271388199595754752259	5, 29*, 59*, 233*, 1103*, 2089*, 3033169*, 107367629*, 536903681*
59	$2^2 \cdot 59$	3 · 13 · 3669631199697131115349471515855850066846063610806992043010594406764144850458064618893717763545170957 99	3, 2833*, 37171*, 179951*, 1824726041*, 3203431780337*
60	$(-1)\,2^4 \cdot 3^2 \cdot 5^2 \cdot 11 \cdot 13 \cdot 31 \cdot 41 \cdot 61$	17 · 5150748653507910906184399685784998327409517035326267521309286916719929747492298535881132936707768267780328207013 1	7*, 3^2*, 11*, 13*, 17*, 5^2*, 31*, 41*, 61*, 151*, 241*, 331*, 1321*, 61681*, 4562284561*
61	$2^2 \cdot 3 \cdot 61$	11 · 11 · 4963366607926258191253263747599075743872279031106013977030931179315068321410043132903311367809803796856443 1	Not given in Lucas' table
62	$(-1)\,2^3 \cdot 5$	41 · 9587677533424712875077490310754244462057883001329733681955351272935859335443594441363194361026847268909460900 1	3*, 5*, 5581*, 8681*, 49477*, 384773*, 715827883*, 2147483647*

THE CANONICAL FORM OF A ONE-PARAMETER GROUP.*

By Philip Franklin.

1. **Introduction.** That the canonical variables of a one-parameter group of transformations involving two variables may be made to play a fundamental role in the theoretical treatment of the solution of certain ordinary differential equations invariant under such a group is well known. In fact the earliest method of integrating a differential equation making use of its invariance under a group to be found by Lie involved the introduction of canonical variables. However, the use of these variables to simplify the usual proofs of some standard theorems, and the fact that they furnish the most direct way of applying the Lie theory to known cases of solvable differential equations, seems to have escaped attention.

It is these applications which concern us here. We begin by finding the nth extended group, and in Theorem I we give a method for forming the most general differential equation of the nth order invariant under such a group. In the light of this theorem, the construction of the usual tables of differential equations invariant under known groups becomes particularly simple.

We next consider second order differential equations invariant under two groups. The two groups may be thought of as giving rise to a two-parameter group, and such two-parameter groups lead to four canonical forms. For one of these canonical forms the reduction apparently requires the solution of a differential equation,† so that we can not reduce the solution of the differential equation to quadratures by a direct introduction of these variables. As an alternative method of solving the differential equation is known‡ which involves nothing beyond quadratures, the results seem to be contradictory. We explain this paradox by showing that if, in the exceptional case, use is made of the coefficients of the differential equation known to be invariant under the two groups, the variables which reduce them to canonical form may be found by quadratures.

2. **The canonical form of a group.** If we start with the symbol§ for the infinitesimal transformation of a group,

* Presented to the American Mathematical Society, December 29, 1926; received March 16, 1927.

† Lie-Scheffers, Vorlesungen über Differentialgleichungen, Leipzig 1891, pp. 424–425, case 4, or A. Cohen, The Lie theory of one-parameter groups, 1911, p. 163.

‡ Lie-Scheffers, l. c. pp. 457–472; or Cohen, l. c. pp. 165–169.

§ Lie-Scheffers, l. c. pp. 49–56, Cohen, l. c. pp. 8–12 or L. E. Dickson, Differential equations from the group standpoint, Annals of Math., (2), vol. 25 (1924), pp. 293–295.

(2.1) $$Uf \equiv \xi(x, y)\, \partial f/\partial x + \eta(x, y)\, \partial f/\partial y,$$

set up the differential system

(2.2) $$\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)} = \frac{dt}{1},$$

and solve the system, obtaining

(2.3) $$v(x, y) = c, \qquad u(x, y) = c' + t,$$

we may write as the equations of the group in finite form

(2.4) $$u(x, y) = u(x_1, y_1) + t, \qquad v(x, y) = v(x_1, y_1).$$

Consequently if we put

(2.5) $$X = u(x, y), \qquad Y = v(x, y),$$

the group becomes the translation

(2.6) $$X = X_1 + t, \qquad Y = Y_1.$$

We call X and Y a pair of *canonical variables* for the group. The differential invariants of the group (2.6) are seen to be

(2.7) $$Y, \qquad dY/dX, \qquad d^2Y/dX^2, \cdots.$$

Consequently the equations of the nth extended group must be

(2.8) $$\frac{dX}{1} = \frac{dY}{0} = \frac{dY'}{0} = \frac{dY''}{0} = \cdots = \frac{dY^{(n)}}{0}.$$

The nth extended group has the same symbol as the group itself,

(2.9) $$U^n f = \partial f/\partial X.$$

From (2.7) the most general differential equation of the nth order invariant under (2.6) is:

(2.91) $$F(Y, Y', Y'', \cdots, Y^{(n)}) = 0.$$

As a change of variables can have no effect on the property of invariance under a group, we see at once on using (2.5) to transform back to the original variables, that the invariants of the nth extended group formed from (2.1) are

(2.92) $$v, \ dv/du, \ d^2v/du^2, \cdots, d^nv/du^n,$$

so that the most general differential equation of the nth order invariant under the group is

(2.93) $$F(v, \ dv/du, \ d^2v/du^2, \cdots, d^nv/du^n) = 0.$$

This establishes

THEOREM I. *If* $v(x, y) = c$, *and* $u(x, y) = c' + t$ *are solutions of the differential equations* (2.2) *formed for a given group, the most general differential equation of the nth order invariant under this group is*

$$F(v, dv/du, \cdots, d^n v/du^n) = 0.$$

Corollary. The differential equation may be written

$$G(I_0, I_1, \cdots, I_n) = 0,$$

where

$$I_0 = f_0(v), \quad I_{k+1} = \frac{df_{k+1}(I_0, I_1, \cdots, I_k)}{d[au + g_{k+1}(I_0, I_1, \cdots, I_{k+1})]},$$

a being any constant, and the functions being arbitrary except for the requirement that the I's are all independent, i. e. I_k *must contain* $y^{(k)}$.

The equation of the corollary is clearly equivalent to that of the theorem, but is often considerably simpler.

The initial determination of $v(x, y)$ involves the solution of a differential equation. Then $u(x, y)$ is found by a quadrature, and the remaining invariants by differentiation alone. That the invariants depend on quadratures alone, after the first is determined, is well known.*

3. **Applications.** We shall now apply our theorem to the specific computation of the invariants for a few groups of general form. By specializing the arbitrary constants and functions in these forms, all the types of the usual tables for differential equations of the first order are obtained.

I. Let $Uf = A(x)\, \partial f/\partial x + B(x)\, C(y)\, \partial f/\partial y$. Here

$$I_0 = v(x, y) = \int \frac{B(x)}{A(x)}\, dx - \int \frac{dy}{C(y)}; \qquad u(x, y) = \int \frac{dx}{A(x)},$$

$$I_1 = dv/du = B(x) - \frac{A(x)}{C(y)}\, y'.$$

The first order differential equation is $I_1 = F(I_0)$. We note the following special cases:

If $A(x) = s$, $B(x) = 1$, $C(y) = -r$, we find

$$1 + \frac{s}{r}\, y' = F\left(\frac{x}{s} + \frac{y}{r}\right) \tag{3.11}$$

or

$$y' = f(rx + sy). \tag{3.1}$$

* Cohen, l. c. p. 52 and p. 201; Lie-Scheffers, l. c. pp. 373–377.

If $A(x) = sx$, $B(x) = 1$, $C(y) = -ry$, we find

$$1+\frac{sx}{ry}y' = F\left(\frac{\log x}{s}+\frac{\log y}{r}\right) \tag{3.21}$$

or

$$xy' = yf(x^r y^s). \tag{3.2}$$

If $A(x) = x^{n+1}/g(x)$, $B(x) = x^n/g(x)$, $C(y) = ny$, we find

$$\frac{x^n}{g(x)}\left(1-\frac{xy'}{ny}\right) = F\left(\log x-\frac{\log y}{n}\right), \tag{3.31}$$

or

$$xy'-ny = g(x)f(y/x^n). \tag{3.3}$$

If $A(x) = 1/g(x)$, $B(x) = x/g(x)$, $C(y) = \mp 1/y$, we find

$$\frac{1}{g(x)}(x\pm yy') = F(x^2\pm y^2). \tag{3.41}$$

If $A(x) = x^{1-r}$, $B(x) = 1$, $C(y) = \mp y^{1-s}$, we find

$$1\pm\frac{x^{1-r}}{y^{1-s}}y' = F\left(\frac{x^r}{r}\pm\frac{y^s}{s}\right) \tag{3.51}$$

or

$$x^{1-r}y^{s-1}y' = f(x^r/r\pm y^s/s). \tag{3.5}$$

II. Let $Uf = B(x)C(y)\,\partial f/\partial y$. Here

$$I_0 = v(x,y) = x;\quad u(x,y) = \frac{1}{B(x)}\int\frac{dy}{C(y)},$$

$$I_1 = dv/du = \left\{-\frac{B'(x)}{[B(x)]^2}\int\frac{dy}{C(y)}+\frac{y'}{B(x)C(y)}\right\}^{-1}.$$

We note in particular the case in which $B(x) = e^{(s-1)\int R(x)\,dx}$, $C(y) = y^s$, where we find

$$e^{(1-s)\int R(x)\,dx}\left[\frac{y'}{y^s}+R(x)y^{1-s}\right] = F(x), \tag{3.61}$$

or

$$y'+R(x)y = Q(x)y^s. \tag{3.6}$$

III. Let

$$Uf = \frac{P_y\,\partial f/\partial x - P_x\,\partial f/\partial y}{P_y Q_x - P_x Q_y}E(Q),$$

where $P(x, y)$ and $Q(x, y)$ are arbitrary functions of two variables, and the subscripts denote partial differentiation. Here

$$I_0 = v(x, y) = P(x, y);\ u(x, y) = \int \frac{Q_x\,dx + Q_y\,dy}{E(Q)},$$

$$I_1 = dv/du = \frac{E(Q)\,[P_x + P_y y']}{Q_x + Q_y y'}.$$

We note two particular cases.

If $P(x, y) = \log y - n \log x$, $Q(x, y) = r \log x + s \log y$, $E(Q) = e^Q$, we find

$$x^r y^s \frac{\frac{y'}{y} - \frac{n}{x}}{\frac{r}{x} + \frac{sy'}{y}} = F(\log y - n \log x), \tag{3.71}$$

or*

$$\frac{x^r y^s (xy' - ny)}{sxy' + ry} = f(y/x^n). \tag{3.7}$$

If $P(x, y) = x^2 - ny^2$, $Q(x, y) = \tan^{-1}(\sqrt{-n}\ y/x)$, $E(Q) = 1$, we find

$$\frac{2(x - nyy')(x^2 - ny^2)}{\sqrt{-n}\,(xy' - y)} = F(x^2 - ny^2), \tag{3.81}$$

or

$$\frac{y - xy'}{x - nyy'} = f(x^2 - ny^2). \tag{3.8}$$

The equations (3.1) to (3.8) written above may be looked on as a list of the more common forms of solvable differential equations of the first order. They are, in fact, identical with the set given by Professor Dickson,† which include those given by earlier writers as special cases. We notice that to integrate these equations we need make no further reference to the theory of groups, since after going back to the forms (3.11) to (3.81), the method of rendering them exact is obvious.

For the practical problem of integrating differential equations, a table giving a large number of special cases is, of course, much more useful than our general cases I, II, and III. In fact, a form as general as III (which includes I as a special case) is little better for direct application than the form $du = 0$, u being the solution we are trying to find! How-

* Both in Cohen, l. c. p. 61, and in Dickson, l. c. p. 324, the sign in the denominator before ry is incorrectly given as minus.

† l. c. p. 324.

ever, by using these forms the construction of more extensive tables is materially simplified.*

4. **Higher order equations.** The integration of a differential equation of order higher than the first known to be invariant under a given one-parameter group may be reduced to the problem of integrating an equation of order lower by unity. By our Theorem I, the setting up of the most general differential equation of the nth order invariant under any of the groups treated in the previous section merely involves further differentiations of the invariants already obtained. The invariants so obtained are, of course, simplified by the application of the corollary. This will be illustrated in a particular case in the next section. As the method of construction of tables for such equations† on the basis of section 3 is obvious, we shall not describe it in detail.

5. **Second order equations invariant under two groups.** The integration of a second order differential equation invariant under two distinct given groups may always be reduced to quadratures. The direct application of the method of canonical variables apparently requires the solution of a differential equation, not in general reducible to quadratures. We shall now explain this paradox. We recall that if two infinitesimal transformations are known leaving a differential equation of the second order invariant, we may find two new infinitesimal transformations, with symbols U_1 and U_2 having the same property, and such that

$$(U_1 U_2)f \equiv 0 \text{ or } U_1 f. \tag{5.11}$$

The two further possibilities,

$$U_2 f \equiv \varrho(x, y)\, U_1 f \text{ or } U_2 f \not\equiv \varrho(x, y)\, U_1 f, \tag{5.12}$$

lead to four cases. Each of these has a corresponding simple canonical form, and in three of these cases the reduction to this form may be accomplished by quadratures alone. The exceptional case is that for which

$$(U_1 U_2)f \equiv U_1 f \text{ and } U_2 f \equiv \varrho(x, y)\, U_1 f. \tag{5.2}$$

That we can *not* carry out the reduction by quadratures alone if we only know the two infinitesimal transformations represented by $U_1 f$ and $U_2 f$ is clear from the following example:

$$\begin{aligned} U_1 f &= \partial f/\partial x + g(x, y)\, \partial f/\partial y, \\ U_2 f &= x\, \partial f/\partial x + x g(x, y)\, \partial f/\partial y. \end{aligned} \tag{5.21}$$

* Compare the amount of computation here given with that used by Cohen, l. c. pp. 52–60.

† Dickson, l. c. p. 349; Cohen, l. c. p. 91, p. 236.

These two symbols satisfy (5.2), but as the canonical form for (5.2) is

$$U_1 f = \partial f/\partial Y, \qquad U_2 f = Y\,\partial f/\partial Y, \tag{5.22}$$

the canonical variable X for (5.21) satisfies:

$$\partial X/\partial x + g(x, y)\,\partial X/\partial y = 0, \tag{5.23}$$

so that $X(x, y) = c$ satisfies

$$dy/dx = g(x, y), \tag{5.24}$$

which can certainly not be solved by quadratures alone for all values of $g(x, y)$.*

If we know not only the two infinitesimal transformations satisfying (5.2), but also a differential equation of the second order invariant under both of them, then, using the coefficients of the given second order equation, we can determine the canonical variables by quadratures. To see this in detail, let us consider two infinitesimal transformations, satisfying (5.2). By taking $\varrho(x, y)$ as a new variable, x_1, we may reduce the symbols to the form:

$$\begin{aligned} U_1 f &= \quad h(x_1, y_1)\,\partial f/\partial x_1 + \quad g(x_1, y_1)\,\partial f/\partial y_1, \\ U_2 f &= x_1 h(x_1, y_1)\,\partial f/\partial x_1 + x_1 g(x_1, y_1)\,\partial f/\partial y_1. \end{aligned} \tag{5.25}$$

We find from these equations and (5.2),

$$(U_1 U_2) f = h^2\,\partial f/\partial x_1 + gh\,\partial f/\partial y_1 \equiv U_1 f, \tag{5.26}$$

and consequently, $h(x_1, y_1) = 1$, and, on dropping subscripts, (5.25) reduces to (5.21). We shall now set up the most general second order differential equation invariant under the two infinitesimal transformations of (5.21).

We first consider the group with symbol

$$U_1 f = \partial f/\partial x + g(x, y)\,\partial f/\partial y. \tag{5.3}$$

* This shows that Theorem 3 (Dickson, l. c. p. 363) is false. The error in Professor Dickson's proof occurs on p. 361, l. 20, where he uses our equation (5.11) to show that the differential equations for the path curves of $U_1 f$ are invariant under $U_2 f$, and concludes that in consequence these equations may be solved by quadratures, overlooking the fact that when the first of our equations (5.12) holds, the group given by $U_2 f$ is *trivial* for $U_1 f$, and hence is of no help towards the integration. Thus the criticism of Lie-Scheffers in the footnotes of p. 363 and p. 365 is unjustified, and Professor Dickson's proof of the theorem on p. 365 needs some such supplement as the discussion of Theorem II which we give here.

To determine the invariants, we have:

$$\frac{dx}{1} = \frac{dy}{g(x, y)} = \frac{dt}{1}. \tag{5.31}$$

Thus, if $v(x, y)$ is a solution of the first equation, so that

$$\partial v/\partial x + g(x, y)\, \partial v/\partial y = 0, \tag{5.32}$$

we have $v(x, y) = c$ and $x = c' + t$ as the first integrals to be used in Theorem I, so that the most general second order differential equation invariant under (5.3) is

$$d^2v/dx^2 = G_1(v, dv/dx). \tag{5.33}$$

For the second group, we have

$$U_2 f = x\, \partial f/\partial x + x g(x, y)\, \partial f/\partial y. \tag{5.34}$$

To find the invariants we solve

$$\frac{dx}{x} = \frac{dy}{x g(x, y)} = \frac{dt}{1}. \tag{5.35}$$

The first integrals are $v(x, y) = c$ and $\log x = c' + t$, so that by Theorem I the most general second order differential equation invariant under (5.34) is

$$F(v,\ x\, dv/dx,\ x^2 d^2v/dx^2) = 0, \tag{5.36}$$

or

$$d^2v/dx^2 = (dv/dx)^2\, G_2(v, x\, dv/dx). \tag{5.37}$$

Comparison of (5.33) and (5.37) shows that the only equations invariant under (5.3) and (5.33) are those of the form

$$d^2v/dx^2 = (dv/dx)^2\, G(v). \tag{5.4}$$

To see what this becomes in terms of x and y, we put:

$$v' = v_x + v_y y', \qquad v'' = v_{xx} + 2 v_{xy} y' + v_{yy} y'^2 + v_y y'', \tag{5.41}$$

obtaining:

$$y'' = \frac{v_x^2 G - v_{xx}}{v_y} + \frac{2(v_x v_y G - v_{xy})\, y'}{v_y} + \frac{(v_y^2 G - v_{yy})\, y'^2}{v_y}. \tag{5.42}$$

As the variables which converted the two given infinitesimal transformations to the form (5.21) must convert the given differential equation to (5.42), on introducing these variables and solving for y'', we shall have:

$$y'' = A(x, y) + 2B(x, y)\, y' + C(x, y)\, y'^2. \tag{5.43}$$

Let us equate coefficients in the last two equations, and make certain reductions by (5.32). We find

$$\begin{aligned}
A &= -g v_x G + g v_{xx}/v_x = g\, \partial/\partial x \Big(\quad \log v_x - \int G\, dv \Big), \\
B &= \quad v_x G - \quad v_{xy}/v_y = \quad \partial/\partial x \Big(-\log v_y + \int G\, dv \Big), \\
B &= -g v_y G + g v_{xy}/v_x = g\, \partial/\partial y \Big(\quad \log v_x - \int G\, dv \Big), \\
C &= \quad v_y G - \quad v_{yy}/v_y = \quad \partial/\partial y \Big(-\log v_y + \int G\, dv \Big).
\end{aligned} \tag{5.44}$$

Consequently, if we form the expression:

$$\int e^{\int \frac{A dx + B dy}{g}} dx + e^{-\int B dx + C dy}\, dy = V(x, y), \tag{5.45}$$

the constants of integration may be so taken that all the functions of two variables to be integrated are exact differentials, and the value of the function will be:

$$V(x, y) = \int C_1\, e^{-\int G dv}\, dv = V(v). \tag{5.46}$$

Thus $V(x, y)$ may be taken as the first integral of (5.31) or (5.35) since it satisfies (5.32). This establishes the following theorem:

THEOREM II. *If a differential equation of the second order invariant under two groups satisfying* (5.2) *is known, the canonical variables for these groups, and hence the reduction to the standard form* (5.22) *may be found by quadratures.*

Specifically, we have merely to introduce the variables

$$x_1 = \varrho(x, y), \qquad y_1 = y, \tag{5.47}$$

determine $V(x_1, y_1)$ by (5.45) and put

$$x_2 = V(x_1, y_1), \qquad y_2 = x_1. \tag{5.48}$$

In terms of the variables of (5.47) the symbols for the groups take the form (5.21), and in terms of the variables of (5.48), the form (5.22).

The integration of the differential equation is at once found from (5.4), and is

$$x = \int C_1 e^{-\int G\, dv}. \tag{5.5}$$

Comparison of this with (5.45) and (5.46) shows that the transformation (5.48) is unnecessary if we merely desire to solve the given differential equation. For, after using (5.47), the equation

$$x_1 = \int e^{\int \frac{A\,dx_1 + B\,dy_1}{g}} dx_1 + e^{-\int B\,dx_1 + C\,dy_1} dy_1, \tag{5.51}$$

with the restriction that the differentials must be made exact, contains two arbitrary constants and gives the general solution of (5.43), and hence on using the inverse of (5.47) the solution of the original equation.

Thus Theorem II is not really necessary if the integration of the differential equation is our sole object. However it rounds out a uniform treatment, by making it possible to treat all cases of integration of a second order differential equation invariant under two groups by means of the canonical variables.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASSACHUSETTS.

A CLASS OF REAL QUADRATIC FORMS IN INFINITELY MANY VARIABLES.*

By F. H. Murray.

In this paper a discussion is given of certain quadratic forms in infinitely many variables, $\sum k_{rs} x_r x_s$, in which k_{rs} is real and such that, for every r, $\sum_s k_{rs}^2$ converges. Such a form is not bounded in general, under the condition $\sum x_s^2 = 1$; it was shown by Carleman† that the theory of these forms may be brought under the theory of integral equations, but the possible applications of the theory to the boundary problems of mathematical physics made it desirable to develop the theory in as direct a manner as possible.

In the first four sections the spectral form corresponding to a given real quadratic form is discussed, with certain properties of related linear forms; in the fifth and sixth sections, respectively, a representation is given of the solutions of an infinite system of ordinary linear equations, and of an infinite system of differential equations of the second order, with constant coefficients. In the remaining sections the preceding theory is applied to the "formal" solution of certain boundary problems for a system of partial differential equations.

1. **Two theorems.** In the following developments two theorems will be used constantly, of which the first is the following:‡

Theorem I: *If $f(x)$ is continuous in the interval $a \leqq x \leqq b$, while $g(x, n)$ has its total variation in this interval less than a constant B for all positive integral values of n, and for every x $\lim_{n\to\infty} g(x, n) = g(x)$, then*

$$\lim_{n\to\infty} \int_a^b f(x)\, d_x\, g(x, n) = \int_a^b f(x)\, d_x\, g(x).$$

Theorem II. *If $u_n^{(r)}$ $v_n^{(r)}$ are real constants satisfying the conditions $|u_n^{(r)}| \leqq w_n$, $\sum_{n=1}^{\infty} |v_n^{(r)}|^2 < C$, $\sum_{n=1}^{\infty} w_n^2$ converges, and $\lim_{r\to\infty} u_n^{(r)} = u_n$, $\lim_{r\to\infty} v_n^{(r)} = v_n$, then the series $\sum_n u_n v_n$ converges, and*

$$\lim_{r\to\infty} \sum_{n=1}^{\infty} u_n^{(r)} v_n^{(r)} = \sum_{n=1}^{\infty} u_n v_n.$$

* Received April 11, 1927, in revised form, August 31, 1927.

† Sur les équations intégrales singulières à noyau réel et symétrique, Uppsala, 1923, also Uppsala Universitets Årsskrift, Matematik och Naturwetenskap, 1923.

‡ Helly, Wien. Sitzungsber. d. Akademie der Wiss., Bd. 121 (1912); Carleman, l. c., p. 9.

To prove the convergence of the series, suppose ε given, and M so chosen that $\sum_{n=m}^{\infty} w_n^2 < \varepsilon$, $m > M$; then from Bessel's inequality

$$\left| \sum_{n=m}^{m+p} u_n^{(r)} v_n^{(r)} \right| \leqq \sqrt{\sum_{n=m}^{m+p} w_n^2} \sqrt{\sum_{n=m}^{m+p} |v_n^{(r)}|^2} \leqq \sqrt{\varepsilon C},$$

whence

$$\left| \sum_{n=m}^{m+p} u_n v_n \right| \leqq \sqrt{\varepsilon C}, \qquad m > M,$$

and the convergence follows from a theorem on limits. Now

$$\left| \sum_{n=1}^{\infty} u_n^{(r)} v_n^{(r)} - \sum_{n=1}^{\infty} u_n v_n \right| \leqq \left| \sum_{n=1}^{N} u_n^{(r)} v_n^{(r)} - \sum_{n=1}^{N} u_n v_n \right| + \left| \sum_{N+1}^{\infty} u_n^{(r)} v_n^{(r)} \right| + \left| \sum_{N+1}^{\infty} u_n v_n \right|$$

if $\varepsilon > 0$ is given, suppose N such that

$$\left| \sum_{N+1}^{\infty} u_n^{(r)} v_n^{(r)} \right| < \frac{\varepsilon}{3}, \qquad \left| \sum_{N+1}^{\infty} u_n v_n \right| < \frac{\varepsilon}{3}.$$

Then R can be so chosen that

$$\left| \sum_{n=1}^{N} u_n^{(r)} v_n^{(r)} - \sum_{n=1}^{N} u_n v_n \right| < \frac{\varepsilon}{3}, \qquad r > R;$$

hence

$$\left| \sum_{n=1}^{\infty} u_n^{(r)} v_n^{(r)} - \sum_{n=1}^{\infty} u_n v_n \right| < \varepsilon,$$

from which the theorem follows.

2. **Real quadratic forms.** Suppose given a set of real numbers k_{rs} $(r, s = 1, 2, \cdots)$ infinite in number, which satisfy the conditions $k_{rs} = k_{sr}$, $\sum_{s=1}^{\infty} k_{rs}^2$ converges. If n is arbitrary, the quadratic form

$$K_n(x) = \sum_{r,s=1}^{n} k_{rs} x_r x_s$$

can be transformed into a sum of squares

$$K_n(x) = \sum_{i=1}^{n} \lambda_i z_i^2$$

by means of a real orthogonal transformation*

$$z_i = l_{i1} x_1 + l_{i2} x_2 + \cdots + l_{in} x_n = L_i^{(n)}(x)$$

* A simple proof of this theorem can be given as follows. Suppose λ a root of the characteristic equation $|k_{rs} - \delta_s^r \lambda| = 0$; and let $l_1', \cdots, l_n'$ be a solution of the system of equations:

the coefficients of which satisfy the conditions

$$\sum_{i=1}^{n} l_{ir}\, l_{is} = \sum_{i=1}^{n} l_{ri}\, l_{si} = \delta_s^r = \begin{matrix} 0, & r \neq s \\ 1, & r = s. \end{matrix}$$

The constants λ_i are real, and some may be zero; in any case,

$$\sum_{i=1}^{n} L_i^{(n)}(x)^2 = \sum_{i=1}^{n} x_i^2 = (x, x)_n, \qquad K_n(x) = \sum_{i=1}^{n} \lambda_i\, L_i^{(n)}(x)^2.$$

From the quadratic form is obtained the bilinear form

$$K_n(x, y) = \sum_{r,s=1}^{n} k_{rs}\, x_r\, y_s = \frac{1}{2} \sum_{i=1}^{n} y_i \frac{\partial K_n}{\partial x_i} = \sum_{i=1}^{n} \lambda_i\, L_i^{(n)}(x)\, L_i^{(n)}(y).$$

If $y_s = \delta_k^s$, $x_r = l_{mr}$, $\sum_{r=1}^{n} k_{kr}\, l_{mr} = \lambda_m\, l_{mk}$, $k, m = 1, \cdots, n$.

3. **The spectral form** $H(xy; \lambda)$. The construction which follows has much in common with the constructions of Hilbert and Carleman.* Suppose

---

S_1 $$\sum_{s=1}^{n} k_{rs}\, l_s^r = \lambda\, l_r^r, \qquad r = 1, \cdots, n, \quad \sum_{s=1}^{n} l_s^{r\,2} = 1.$$

Suppose the coefficients $(l_1^m, l_2^m, \cdots, l_n^m)$ determined by solving in succession the systems

S_2 $$\sum_{s=1}^{n} l_s^1\, l_s^2 = 0, \qquad \sum_{s=1}^{n} l_s^{2\,2} = 1,$$

S_3 $$\sum_{s=1}^{n} l_s^1\, l_s^3 = 0, \qquad \sum_{s=1}^{n} l_s^2\, l_s^3 = 0, \qquad \sum_{s=1}^{n} l_s^{3\,2} = 1,$$

. . . . . .

S_n $$\sum_{s=1}^{n} l_s^r\, l_s^n = 0, \qquad r = 1, \cdots, n-1, \qquad \sum_{s=1}^{n} l_s^{n\,2} = 1.$$

From these equations it follows that the transformation $x_i = \sum_{r=1}^{n} l_i^r\, z_r$, $(i = 1, \cdots, n)$ satisfies the condition $\sum x_i^2 = \sum z_i^2$, hence $\sum_i l_r^i\, l_s^i = \delta_s^r$ and the inverse transformation becomes $z_r = \sum_i l_i^r\, x_i$. Substituting,

$$\sum_{r,s} k_{rs}\, x_r\, x_s = \sum_{r,s,i,j} k_{rs}\, l_r^i\, l_s^j\, z_i\, z_j = \sum_{r,s} k_{rs}\, l_r^1\, l_s^1\, z_1^2 + \sum_{r,s} \sum_{j=2}^{n} k_{rs}\, l_r^1\, l_s^j\, z_1\, z_s + M(z_2, \cdots, z_n)$$
$$= \lambda\, z_1^2 + M(z_2, \cdots, z_n),$$

and the theorem follows by induction.

From the theory of real quadratic forms it is known that the constants λ_k are necessarily real, hence the coefficients l_{rs} of the transformation can be assumed real, as they satisfy systems of equations with real coefficients.

* Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, pp. 109–125; Carleman, l. c.

$$h_{rs}^{(n)}(\lambda) = 0, \qquad \lambda < \lambda_k^{(n)}, \qquad k = 1, \cdots, n,$$
$$h_{rs}^{(n)}(\lambda) = \sum_{\lambda_k < \lambda} l_{kr}^{(n)}\, l_{ks}^{(n)}, \qquad r \leqq n,\ s \leqq n,$$
$$h_{rs}^{(n)}(\lambda) = 0, \qquad r > n,\ s > n.$$

If $\Delta_1, \Delta_2, \cdots, \Delta_m$ is any set of consecutive intervals, and

$$\Delta_i\, h_{rs}^{(n)} = \sum_{\lambda_{i-1}' < \lambda_k \leqq \lambda_i'} l_{kr}^{(n)}\, l_{ks}^{(n)}, \qquad \lambda_0 = \lambda_1',\ \lambda_m = \lambda_2'$$

then

$$\sum_{i=1}^{m} |\Delta_i\, h_{rs}^{(n)}| \leqq \sum_{\lambda_1' < \lambda_k^{(n)} \leqq \lambda_2'} |l_{kr}^{(n)}\, l_{ks}^{(n)}| \leqq \sqrt{\left(\sum_{k=1}^{n} l_{kr}^{(n)2}\right)\left(\sum_{k=1}^{n} l_{ks}^{(n)2}\right)} = 1,$$

and we may write

$$h_{rs}^{(n)}(\lambda) = \overset{+}{h}{}_{rs}^{(n)}(\lambda) - \overset{-}{h}{}_{rs}^{(n)}(\lambda), \qquad 1 \geqq \overset{+}{h}{}_{rs}^{(n)} + \overset{-}{h}{}_{rs}^{(n)}$$

where $\overset{+}{h}{}_{rs}^{(n)}$, $\overset{-}{h}{}_{rs}^{(n)}$ are nondecreasing positive functions of λ, since $h_{rs}^{(n)}(\lambda)$ vanishes for $\lambda = -\infty$.

Let the set of all rational values of λ be arranged in a sequence $\lambda', \lambda'', \cdots \lambda^{(r)}, \cdots$; since $\overset{+}{h}{}_{11}^{(n)}(\lambda)$ is bounded for all values of n, a subsequence can be found for which this function approaches a limit $\overset{+}{h}{}_{11}(\lambda)$ for every rational λ. Since the sets $(r, s, +)$, $(r, s, -)$ form a denumerable set, it follows that the functions $\overset{+}{h}{}_{rs}^{(n)}(\lambda)$, $\overset{-}{h}{}_{rs}^{(n)}(\lambda)$ can be arranged in a sequence $h_1^{(n)}(\lambda), h_2^{(n)}(\lambda), \cdots, h_m^{(n)}(\lambda), \cdots, h_1^{(n)}(\lambda) = \overset{+}{h}{}_{11}^{(n)}(\lambda)$. To define a limit of $h_{k+1}^{(n)}(\lambda)$ as n becomes infinite, choose a subsequence of the sequence for which $h_k^{(n)}(\lambda)$ approaches a limit for rational λ. If $(n_1^{(k)}, n_2^{(k)}, \cdots, n_k^{(k)}, \cdots)$ is the sequence for which $h_k^{(n)}$ approaches a limit, the sequence $(n_1^{(1)}, n_2^{(2)}, \cdots, n_k^{(k)}, \cdots)$ is such that for every r, its numbers belong to the sequence for which $h_r^{(n)}(\lambda)$ approaches a limit for rational λ, $k > r$. Let the limits be

$$\lim_{n_k \to \infty} \overset{+}{h}{}_{rs}^{(n_k)} = \overset{+}{h}{}_{rs}(\lambda), \qquad \lim_{n_k \to \infty} \overset{-}{h}{}_{rs}^{(n_k)} = \overset{-}{h}{}_{rs}(\lambda), \qquad \lim_{n_k \to \infty} h_{rs}^{(n_k)} = h_{rs}(\lambda).$$

The sequence defined above defines a unique limit for $\overset{+}{h}{}_{rs}^{(n)}$, $\overset{-}{h}{}_{rs}^{(n)}$ at every point λ such that $\overset{+}{h}{}_{rs}(\lambda + 0) = \overset{+}{h}{}_{rs}(\lambda - 0)$, $\overset{-}{h}{}_{rs}(\lambda + 0) = \overset{-}{h}{}_{rs}(\lambda - 0)$. For suppose λ any such value, and choose rational values $\lambda, \lambda_i, \lambda_j$ such that $\lambda_i < \lambda < \lambda_j$, $\lim_{i \to \infty} \lambda_i = \lim_{j \to \infty} \lambda_j = \lambda$. For every n,

$$\overset{+}{h}{}_{rs}^{(n)}(\lambda_i) \leqq \overset{+}{h}{}_{rs}^{(n)}(\lambda) \leqq \overset{+}{h}{}_{rs}^{(n)}(\lambda_j), \qquad \overset{-}{h}{}_{rs}^{(n)}(\lambda_i) \leqq \overset{-}{h}{}_{rs}^{(n)}(\lambda) \leqq \overset{-}{h}{}_{rs}^{(n)}(\lambda_j).$$

Hence if $\overset{+}{h}'_{rs}, \overset{-}{h}'_{rs}$ are the upper limits of the functions $\overset{+}{h}^{(n)}_{rs}(\lambda), \overset{-}{h}^{(n)}_{rs}(\lambda)$,

$$\overset{+}{h}_{rs}(\lambda_i) \leqq \overset{+}{h}'_{rs}(\lambda) \leqq \overset{+}{h}_{rs}(\lambda_j), \qquad \overset{-}{h}_{rs}(\lambda_i) \leqq \overset{-}{h}'_{rs}(\lambda) \leqq \overset{-}{h}_{rs}(\lambda_j).$$

These inequalities must hold for the lower limits also; allowing i and j to become infinite, we obtain the results

$$\overset{+}{h}'_{rs}(\lambda) = \overset{+}{h}_{rs}(\lambda \pm 0), \qquad \overset{-}{h}'_{rs}(\lambda) = \overset{-}{h}_{rs}(\lambda \pm 0).$$

The same equations hold for the lower limits, respectively, consequently the upper and lower limits coincide.

From the sequence which converges at rational values of λ can be obtained a subsequence which converges at the points of discontinuity of the functions $\overset{+}{h}_{rs}, \overset{-}{h}_{rs}$ as these points form a denumerable set; the sequence so obtained is such that the functions $\overset{+}{h}^{(n)}_{rs}, \overset{-}{h}^{(n)}_{rs}$ converge to limiting values for all values of λ.

In what follows n will be understood to refer to a term of the final sequence defined above.

4. **Properties of the functions $h_{rs}(\lambda)$.** From the definition,

$$\sum_{s=1}^{\infty} h^{(n)2}_{rs} = \sum_{s} \sum_{\lambda_k<\lambda} \sum_{\lambda_{k'}<\lambda} l^{(n)}_{kr}\, l^{(n)}_{ks}\, l^{(n)}_{k'r}\, l^{(n)}_{k's} = \sum_{k,k'} l^{(n)}_{kr}\, l^{(n)}_{k'r} \sum_{s=1}^{\infty} l^{(n)}_{ks}\, l^{(n)}_{k's} = \sum_{\lambda_k<\lambda} l^{(n)2}_{kr} \leqq 1.$$

Let $f^{(n)}_r(x, \lambda) = \sum_s h^{(n)}_{rs}(\lambda) x_s$, $f_r(x, \lambda) = \sum_{s=1}^{\infty} h_{sr}(\lambda) x_s$, then from Theorem I,

$$\lim_{n\to\infty} \sum_s h^{(n)}_{rs} x_s = \sum_{s=1}^{\infty} h_{rs}(\lambda) x_s, \qquad \lim_{n\to\infty} f^{(n)}_r(x, \lambda) = f_r(x, \lambda).$$

Again,

$$\sum_r [f^{(n)}_r(x, \lambda)]^2 = \sum_{r=1}^{n} \sum_{s,t} h^{(n)}_{rs}\, h^{(n)}_{rt}\, x_s x_t = \sum_{\lambda_k, \lambda_{k'}<\lambda} \sum_{s,t} \left[\sum_{r=1}^{n} l^{(n)}_{kr}\, l^{(n)}_{ks}\, l^{(n)}_{k'r}\, l^{(n)}_{k's}\right] x_s x_t$$

$$= \sum_{s,t} \left(\sum_{\lambda_k<\lambda} l^{(n)}_{ks}\, l^{(n)}_{kt}\right) x_s x_t = \sum_{\lambda_k<\lambda} L^{(n)}_k(x)^2 \leqq (x, x)_n \leqq (x, x).$$

Consequently, if $\sum_{s=1}^{\infty} x_s^2$, $\sum_{s=1}^{\infty} y_s^2$ converge, $H^{(n)}(x, y, \lambda) = \sum_r f^{(n)}_r(x, \lambda) y_r$, then from the same theorem,

$$\lim_{n\to\infty} H^{(n)}(x, y, \lambda) = \sum_{r=1}^{\infty} f_r(x, \lambda) y_r = H(x, y, \lambda).$$

Now if $\Delta_1, \Delta_2, \cdots, \Delta_m$ is any set of consecutive intervals, with λ_1, λ_2 as end-points,

$$\sum_{i=1}^{m} |\Delta_i f^{(n)}_r(x)| \leqq \sum_{\lambda_1 \leqq \lambda_k < \lambda_2} |l^{(n)}_{kr}\, L^{(n)}_k(x)| \leqq \sqrt{\sum_k l^{(n)2}_{kr}} \sqrt{\sum_k L^{(n)}_k(x)^2} \leqq \sqrt{(x, x)_n};$$

hence $f_r^{(n)}(x, \lambda)$ is of bounded variation, with its total variation independent of n. The function $f_r(x, \lambda)$ must therefore have its total variation less than or equal to $V(x, x)$.

From the preceding inequalities it is easily seen that the series $\sum_s h_{rs}^2$ converges. For if m is arbitrary, $\sum_{s=1}^{m} h_{rs}^{(n)2} \leqq 1$, consequently $\sum_{s=1}^{m} h_{rs}^2 \leqq 1$ and the given series has its partial sums $\leqq 1$; as the partial sum increases with the number of terms, it approaches a limit which is the sum. A similar argument shows that the series $\sum_r f_r(x, \lambda)^2$ also converges.

Fom the definition of these functions it is known that $h_{rs}^{(n)}(\infty) - h_{rs}^{(n)}(-\infty) = \delta_s^r$; a similar equation for the limiting functions is obtained as follows. From the fundamental equations

$$\sum_{i=1}^{n} k_{ri}\, l_{ki}^{(n)} = \lambda_k\, l_{kr}^{(n)} \qquad r = 1, \cdots, n, \quad k = 1, \cdots, n \tag{4.1}$$

is obtained

$$\sum_{i=1}^{n} k_{ri} \left(\sum_{\lambda_1 \leqq \lambda_k < \lambda_2} \frac{l_{ki}^{(n)}\, l_{ks}^{(n)}}{\lambda_k} \right) = \sum_{\lambda_1 \leqq \lambda_k < \lambda_2} l_{kr}^{(n)}\, l_{ks}^{(n)} = \Delta h_{rs}^{(n)},$$

if λ_1, λ_2 are both positive or both negative. Replacing the sum by the corresponding Stieltjes integral, let

$$z_{is}^{(n)} = \sum_{\lambda_1 \leqq \lambda_k < \lambda_2} \frac{l_{ki}^{(n)}\, l_{ks}^{(n)}}{\lambda_k} = \int_{\Delta} \frac{1}{\lambda}\, d\lambda\, h_{is}^{(n)}(\lambda);$$

then

$$\sum_{i=1}^{n} z_{is}^{(n)2} = \sum_{k,k'} \Delta \sum_i \frac{l_{ki}^{(n)}\, l_{ks}^{(n)}\, l_{k'i}^{(n)}\, l_{k's}^{(n)}}{\lambda_k\, \lambda_{k'}} = \sum_{\lambda_1 \leqq \lambda_k < \lambda_2} \frac{l_{ks}^{(n)2}}{\lambda_k^2} \leqq \frac{1}{\lambda_i^2} \sum_k l_{ks}^{(n)2} \leqq \frac{1}{\lambda_i^2}$$

where λ_i is the number λ_1 or λ_2 having the smaller absolute value. Consequently,

$$|\Delta h_{rs}^{(n)}| \leqq \sqrt{\sum_{i=1}^{n} k_{ri}^2}\, \sqrt{\sum_{i=1}^{n} z_{is}^{(n)2}} \leqq \frac{1}{|\lambda_i|} \sqrt{\sum_{i=1}^{\infty} k_{ri}^2}$$

and the same inequality must hold for the limit:

$$|\Delta h_{rs}| \leqq \frac{1}{|\lambda_i|} \sqrt{\sum_{i=1}^{\infty} k_{ri}^2}\,. \tag{4.2}$$

Now

$$[h_{rs}]_{-\infty}^{+\infty} - [h_{rs}^{(n)}]_{-\infty}^{+\infty} = [h_{rs}]_{\lambda}^{\infty} + [h_{rs}]_{-\infty}^{-\lambda} - [h_{rs}^{(n)}]_{\lambda}^{\infty} - [h_{rs}^{(n)}]_{-\infty}^{-\lambda}$$
$$+ [h_{rs}(\lambda) - h_{rs}^{(n)}(\lambda)] - [h_{rs}(-\lambda) - h_{rs}^{(n)}(-\lambda)].$$

Given ε, the sum of the first four terms on the right can be made less than $\varepsilon/2$, by a proper choice of λ, independent of n; then n can be so chosen that each of the remaining brackets on the right is less than $\varepsilon/4$ in absolute value. Consequently

$$\left|[h_{rs}]_{-\infty}^{+\infty} - [h_{rs}^{(n)}]_{-\infty}^{+\infty}\right| < \varepsilon$$

and

$$h_{rs}(\infty) - h_{rs}(-\infty) = \delta_s^r. \tag{4.3}$$

With the same definition of $z_{is}^{(n)}$, assume convergent the series $\sum_s y_s^2$;

$$\sum_{s=1}^{n} z_{is}^{(n)}\, y_s = \sum_{\lambda_1 \leqq \lambda_k < \lambda_2} \frac{l_{ki}^{(n)}\, L_k^{(n)}(y)}{\lambda_k} = \int_\Delta \frac{1}{\lambda}\, d\lambda\, f_i^{(n)}(y, \lambda),$$

$$\left|\sum_{s=1}^{n} z_{js}^{(n)}\, y_s\right| \leqq \frac{1}{|\lambda_i|} \sqrt{\sum_k l_{kj}^{(n)2}} \sqrt{\sum_k L_k^{(n)}(y)^2} \leqq \frac{\sqrt{(y, y)}}{|\lambda_i|}$$

and

$$\left|\int_{\lambda_1}^{\lambda_2} \frac{1}{\lambda}\, d\lambda\, f_j^{(n)}(y, \lambda)\right| \leqq \frac{1}{|\lambda_i|} \sqrt{(y, y)}.$$

From the equations above,

$$\begin{aligned} |\Delta f_r^{(n)}(y)| &= \left|\sum_{s=1}^{n} \Delta h_{rs}^{(n)}\, y_s\right| = \left|\sum_{j=1}^{n} k_{rj} \left(\sum_k \frac{l_{kj}^{(n)}\, L_k^{(n)}(y)}{\lambda_k}\right)\right| \\ &\leqq \sqrt{\sum_{j=1}^{\infty} k_{rj}^2} \sqrt{\sum_{k,k'} \left(\sum_{j=1}^{n} \frac{l_{kj}^{(n)}\, L_k^{(n)}(y)\, l_{k'j}^{(n)}\, L_{k'}^{(n)}(y)}{\lambda_k\, \lambda_{k'}}\right)} \\ &\leqq \frac{1}{|\lambda_i|} \sqrt{\sum_s k_{rs}^2} \sqrt{\sum_k L_k^{(n)}(y)^2} \leqq \frac{\sqrt{\sum_s k_{rs}^2}\, \sqrt{(y, y)}}{|\lambda_i|}. \end{aligned}$$

Hence in the limit,

$$|\Delta f_r(y)| \leqq \frac{1}{|\lambda_i|} \sqrt{\sum_s k_{rs}^2}\, \sqrt{(y, y)}. \tag{4.4}$$

If λ_1, λ_2 are arbitrary, and Δ is the interval $\lambda_1 \leqq \lambda < \lambda_2$, from (4.1),

$$\sum_{\lambda_1 \leqq \lambda_k < \lambda_2} \sum_{i=1}^{n} k_{ri}\, l_{ki}^{(n)}\, L_k^{(n)}(y) = \sum_{i=1}^{n} k_{ri} \sum_s \Delta h_{is}^{(n)}\, y_s = \sum_{\lambda_1 \leqq \lambda_k < \lambda_2} \lambda_k\, l_{kr}^{(n)}\, L_k^{(n)}(y)$$

and this can be written

$$\sum_{i=1}^{n} k_{ri}\, \Delta f_i^{(n)}(y) = \int_{\lambda_1}^{\lambda_2} \lambda\, d\lambda\, f_r^{(n)}(y, \lambda).$$

The limit on the left exists, from Theorem II; applying Theorem I,

$$\sum_{i=1}^{\infty} k_{ri}\,\Delta f_i(y) = \int_{\lambda_1}^{\lambda_2} \lambda\, d\lambda\, f_r(y,\lambda). \tag{4.5}$$

Applying Theorem II again,

$$\sum_{i=1}^{\infty} k_{ri}\,[f_i(y,\infty) - f_i(y,\lambda_1)] = \int_{\lambda_1}^{\infty} \lambda\, d\lambda\, f_r(y,\lambda),$$

$$\sum_{i=1}^{\infty} k_{ri}\,[f_i(y,\lambda_1) - f_i(y,-\infty)] = \int_{-\infty}^{\lambda_1} \lambda\, d\lambda\, f_r(y,\lambda).$$

Hence

$$\sum_{i=1}^{\infty} k_{ri}\,[f_i(y,\lambda)]_{-\infty}^{+\infty} = \int_{-\infty}^{+\infty} \lambda\, d\lambda\, f_r(y,\lambda).$$

Now

$$[f_i(y,\lambda)]_{-\infty}^{+\infty} = \sum_{s=1}^{\infty} y_s\,[h_{is}]_{-\infty}^{+\infty} = y_i$$

hence

$$\int_{-\infty}^{+\infty} \lambda\, d\lambda\, f_r(y,\lambda) = \sum_{i=1}^{\infty} k_{ri}\, y_i. \tag{4.6}$$

In particular,*

$$\int_{-\infty}^{+\infty} \lambda\, d\lambda\, h_{rs}(\lambda) = k_{rs}. \tag{4.7}$$

In (4.5) suppose $y_t = \delta_t^s$; the equation becomes,

$$\sum_{i=1}^{\infty} k_{ri}\,\Delta h_{is} = \int_{\lambda_1}^{\lambda_2} \lambda\, d\lambda\, h_{rs}(\lambda).$$

Suppose λ_ν a point of discontinuity of at least one function $h_{ij}(\lambda)$, and suppose

$$l_{rs}(\lambda_\nu) = h_{rs}(\lambda_\nu + 0) - h_{rs}(\lambda_\nu - 0), \qquad \lambda_1 = \lambda_\nu - \varepsilon,\ \lambda_2 = \lambda_\nu + \varepsilon.$$

Then it can be shown that

$$\sum_{i=1}^{\infty} k_{ri}\, l_{is}(\lambda_\nu) = \lambda_\nu\, l_{rs}(\lambda_\nu). \tag{4.8}$$

For from Theorem II,

$$\lim_{\varepsilon=0} \sum_{i=1}^{\infty} k_{ri}\,\Delta h_{is} = \sum_{i=1}^{\infty} k_{ri}\, l_{is}(\lambda_\nu).$$

Since

$$\int_{\lambda_\nu-\varepsilon}^{\lambda_\nu+\varepsilon} \lambda\, d\lambda\, h_{rs}(\lambda) = \lambda_\nu\,\Delta h_{rs}(\lambda_\nu) + \int_{\lambda_\nu-\varepsilon}^{\lambda_\nu+\varepsilon} (\lambda - \lambda_\nu)\, d\lambda\, h_{rs}(\lambda)$$

* Formulae (4.6) and (4.7) correspond to (66) and (74), respectively, in the work of Carleman, l. c.

and

$$\left| \int_{\lambda_\nu - \varepsilon}^{\lambda_\nu + \varepsilon} (\lambda - \lambda_\nu)\, d_\lambda\, h_{rs}(\lambda) \right| \leqq \varepsilon \int_{\lambda_\nu - \varepsilon}^{\lambda_\nu + \varepsilon} |d_\lambda\, h_{rs}(\lambda)|$$

it follows that

$$\lim_{\varepsilon = 0} \int_{\lambda_\nu - \varepsilon}^{\lambda_\nu + \varepsilon} \lambda\, d_\lambda\, h_{rs}(\lambda) = \lambda_\nu\, l_{rs}(\lambda_\nu)$$

and (4.8) follows. By means of a discussion similar to that given at the beginning of this section it can be shown that $\sum_s \Delta h_{rs}^2 \leqq 1$, consequently $\sum_{s=1}^{\infty} l_{rs}(\lambda_\nu)^2 \leqq 1$.

5. Solution of the linear equations:

$$\sum_{s=1}^{\infty} k_{rs}\, x_s - \lambda x_r = y_r, \qquad r = 1, 2, 3, \cdots \tag{5.1}$$

$$\sum_s y_s^2 \text{ convergent.}$$

Suppose $(x_1, x_2, \cdots, x_n)$ a solution of the system

$$\sum_{s=1}^{n} k_{rs}\, x_s - \lambda x_r = y_r, \qquad r = 1, \cdots, n.$$

From the identity

$$K_n(xy) = \sum_{k=1}^{n} \lambda_k^{(n)}\, L_k^{(n)}(x)\, L_k^{(n)}(y)$$

there results

$$\sum_{s=1}^{n} k_{rs}\, x_s - \lambda x_r = \sum_{k=1}^{n} \lambda_k^{(n)}\, L_k^{(n)}(x)\, l_{kr}^{(n)} - \lambda x_r = y_r,$$

whence

$$L_m^{(n)}(y) = \sum_{k=1}^{n} \lambda_k^{(n)}\, L_k^{(n)}(x) \sum_{r=1}^{n} l_{kr}^{(n)}\, l_{mr}^{(n)} - \lambda\, L_m^{(n)}(x),$$

$$L_m^{(n)}(x) = \frac{L_m^{(n)}(y)}{\lambda_m^{(n)} - \lambda}$$

and finally

$$x_k^{(n)} = \sum_{m=1}^{n} \frac{l_{mk}^{(n)}\, L_m^{(n)}(y)}{\lambda_m^{(n)} - \lambda} = \int_{-\infty}^{+\infty} \frac{1}{\mu - \lambda}\, d_\mu f_k^{(n)}(y, \mu). \tag{5.2}$$

If $\lambda = \alpha + i\beta$,

$$\frac{1}{\mu - \lambda} = \frac{\mu - \alpha}{(\mu - \alpha)^2 + \beta^2} + \frac{i\beta}{(\mu - \alpha)^2 + \beta^2}, \quad \left| \frac{1}{\mu - \lambda} \right| \leqq \frac{1}{|\beta|}, \quad \text{Max} \frac{1}{|\mu - \alpha|}.$$

Assume $\beta \neq 0$; if the real and imaginary parts of the integral in (5.2) are considered separately, $l > |\alpha|$,

$$\left|\int_l^\infty \frac{\mu-\alpha}{(\mu-\alpha)^2+\beta^2}\, d_\mu f_k^{(n)}(y,\mu)\right| \leqq \text{Max}\,\frac{1}{|\mu-\alpha|}\int_l^\infty |d_\mu f_k^{(n)}(y,\mu)| \leqq \frac{(y,y)}{l-\alpha},$$

and similarly,

$$\left|\int_l^\infty \frac{\beta}{(\mu-\alpha)^2+\beta^2}\, d_\mu f_k^{(n)}(y,\mu)\right| \leqq \frac{(y,y)}{l-\alpha}.$$

Now

$$x_k^{(n)} = \int_{-l}^{l}\frac{1}{\mu-\lambda}\, d_\mu f_k^{(n)}(y,\mu) + \left(\int_{-\infty}^{-l}+\int_l^\infty\right)\frac{\mu-\alpha}{(\mu-\alpha)^2+\beta^2}\, d_\mu f_k^{(n)} + i\beta\left(\int_{-\infty}^{-l}+\int_l^\infty\right)\frac{1}{(\mu-\alpha)^2+\beta^2}\, d_\mu f_k^{(n)}$$

$$= I_1 + I_2 + I_3,$$

and from the preceding inequalities and the corresponding ones for $\mu<0$,

$$|I_2+I_3| \leqq \frac{4(y,y)}{l-|\alpha|}.$$

If the integral I_1 is broken up into its real and imaginary parts, from Theorem I

$$\lim_{n\to\infty} I_1 = \lim_{n\to\infty}\int_{-l}^{l}\frac{1}{\mu-\lambda}\, d_\mu f_k^{(n)}(y,\mu) = \int_{-l}^{l}\frac{1}{\mu-\lambda}\, d_\mu f_k(y,\mu).$$

In the same manner as above can be obtained the inequality

$$\left|\left(\int_{-\infty}^{-l}+\int_l^\infty\right)\frac{1}{\mu-\lambda}\, d_\mu f_k(y,\mu)\right| \leqq \frac{4(y,y)}{l-|\alpha|}.$$

Now

$$\left|\int_{-\infty}^{\infty}\frac{1}{\mu-\lambda}\, d_\mu f_k^{(n)}(y,\mu) - \int_{-\infty}^{\infty}\frac{1}{\mu-\lambda}\, d_\mu f_k(y,\mu)\right| \leqq \left|\int_{-l}^{l}\frac{1}{\mu-\lambda}\, d_\mu f_k^{(n)} - \int_{-l}^{l}\frac{1}{\mu-\lambda}\, d_\mu f_k\right| + \frac{8(y,y)}{l-|\alpha|}.$$

The right-hand member can be made arbitrarily small by choosing first l and then n; consequently,

$$\lim_{n\to\infty} x_k^{(n)} = x_k = \int_{-\infty}^{\infty}\frac{1}{\mu-\lambda}\, d_\mu f_k(y,\mu).$$

It is easily shown that the sum of the squares of the absolute values of the x's converges. For if $x_k^{(n)} = z_k' + i z_k''$,

$$z_k' = \sum_{m=1}^{n}\frac{\lambda_m-\alpha}{(\lambda_m-\alpha)^2+\beta^2}\, l_{mk}^{(n)}\, L_m^{(n)}(y), \qquad z_k'' = \sum_{m=1}^{n}\frac{\beta}{(\lambda_m-\alpha)^2+\beta^2}\, l_{mk}^{(n)}\, L_{mk}^{(n)}(y)$$

and

$$\sum_{k=1}^{n} z_k'^2 = \sum_{m,r} \frac{(\lambda_m - \alpha)(\lambda_r - \alpha)}{[(\lambda_m - \alpha)^2 + \beta^2][(\lambda_r - \alpha)^2 + \beta^2]} \sum_{k=1}^{n} (l_{mk}^{(n)} l_{rk}^{(n)}) L_m^{(n)}(y) L_r^{(n)}(y)$$

$$= \sum_m \left[\frac{(\lambda_m - \alpha) L_m^{(n)}(y)}{(\lambda_m - \alpha)^2 + \beta^2}\right]^2 \leqq \frac{(y, y)}{\beta^2}, \quad \leqq (y, y) \operatorname{Max}\left[\frac{1}{\lambda_m - \alpha}\right]^2.$$

Similarly,

$$\sum_{k=1}^{n} z_k''^2 \quad \leqq \frac{(y, y)}{\beta^2}, \quad \leqq (y, y) \operatorname{Max}\left[\frac{1}{\lambda_m - \alpha}\right]^2$$

and

$$\sum_{k=1}^{n} |x_k^{(n)}|^2 \leqq \frac{2(y, y)}{\beta^2}, \quad \leqq 2(y, y) \operatorname{Max}\left[\frac{1}{\lambda_m - \alpha}\right]^2. \tag{5.3}$$

From the last inequality, it is seen that the partial sums $\sum_{s=1}^{N} |x_s|^2$ have an upper bound, and the convergence of the series $\sum_s |x_s|^2$ follows.

To prove that the infinite system of equations is satisfied, suppose for convenience that $x_s^{(n)} = 0$, $s > n$; let

$$S_r^n = \sum_{s=1}^{\infty} k_{rs} x_s^{(n)} = \lambda x_r^{(n)} + y_r, \qquad r \leqq n,$$

then from Theorem II,

$$\lim_{n \to \infty} S_r^n = \sum_{s=1}^{\infty} k_{rs} x_s = \lambda x_r + y_r, \quad r = 1, 2, \cdots, \infty.$$

Hence if the constants k_{rs} satisfy the conditions specified at the beginning of section 2, if λ is not real, and the series $\sum_s y_s^2$ converges, the given system of linear equations possesses at least one solution (x) such that $\sum |x_s^2|$ converges. Several solutions may exist and satisfy this condition; if only one exists the system of equations is of Class 1, as this class is defined by Carleman.

Also, suppose that it is known that the roots $\lambda_k^{(n)}$ are all greater than some positive constant q; then in the integrals representing the solution, $|1/\lambda_m| \leqq 1/q$, and the solution can be written

$$x_k = \int_q^{\infty} \frac{1}{\mu - \lambda} \, d\mu \, f_k(y, \mu).$$

The integrand $1/(\mu - \lambda)$ is bounded when $\lambda = 0$; the existence of the solution follows as before, and the convergence of the sum of the squares of the absolute values follows from the second inequality of (5.3).

6. **An infinite system of differential equations.** By means of the formulae of the preceding paragraphs it is possible to solve the infinite system of differential equations

$$(6.1) \qquad \frac{d^2 x_r}{dt^2} = -\sum_{s=1}^{\infty} k_{rs}\, x_s, \qquad r = 1, 2, \cdots$$

with the initial conditions

$$(6.11) \quad [x_r]_0 = y_r', \qquad \left[\frac{dx_r}{dt}\right]_0 = y_r'', \qquad \sum_s y_r'^2,\ \sum_s y_r''^2 \text{ convergent.}$$

From the equations

$$\sum_{s=1}^{n} k_{rs}\, l_{ks}^{(n)} = \lambda_k\, l_{kr}^{(n)}, \qquad r, k = 1, \cdots, n,$$

the required solution can be obtained by a passage to the limit. Suppose $\omega(\lambda)$ continuous on the interval $\lambda' \leqq \lambda \leqq \lambda''$; from the equations above, with the conditions assumed at the beginning of section 4,

$$\sum_{s=1}^{n} k_{rs}\left[\sum_{\lambda' \leqq \lambda_k < \lambda''} \omega(\lambda_k)\, l_{ks}^{(n)}\, L_k^{(n)}(y)\right] = \sum_{\lambda' \leqq \lambda_k < \lambda''} \lambda_k\, \omega(\lambda_k)\, l_{kr}^{(n)}\, L_k^{(n)}(y).$$

Suppose $y_1, y_2, \cdots$ a set of real numbers such that $\sum_s y_s^2$ converges, and let

$$z_s^{(n)} = \sum_{\lambda' \leqq \lambda_k < \lambda''} \omega(\lambda_k)\, l_{ks}^{(n)}\, L_k^{(n)}(y), \qquad = \int_{\lambda'}^{\lambda''} \omega(\lambda)\, d_\lambda f_s^{(n)}(y, \lambda),$$

$$z_s = \int_{\lambda'}^{\lambda''} \omega(\lambda)\, d_\lambda f_s(y, \lambda).$$

Then by means of a reduction already employed,

$$(6.2) \qquad \sum_{s=1}^{n} z_s^{(n)2} = \sum_{\lambda' \leqq \lambda_k < \lambda''} \omega(\lambda_k)^2\, L_k^{(n)}(y)^2 \leqq (y, y)_n \operatorname{Max} \omega(\lambda_k)^2.$$

Substituting in the equation above, we obtain

$$\sum_{s=1}^{n} k_{rs}\, z_s^{(n)} = \int_{\lambda'}^{\lambda''} \lambda\, \omega(\lambda)\, d_\lambda f_r^{(n)}(y, \lambda).$$

From Theorems I, II,

$$(6.2) \qquad \sum_{s=1}^{\infty} k_{rs}\, z_s = \int_{\lambda'}^{\lambda''} \lambda\, \omega(\lambda)\, d_\lambda f_r(y, \lambda).$$

In what follows it will be assumed that the roots $\lambda_k^{(n)}$ are all greater than some constant $q > 0$. In (6.2) let $\omega(\lambda) = \cos\sqrt{\lambda}\, t$; suppose $\lambda' = 0$, and let λ'' become infinite. From Theorems II, I,

$$\lim_{\lambda'' \to \infty} \sum_{s=1}^{\infty} k_{rs}\, z_s = \sum_{s=1}^{\infty} k_{rs} \int_0^{\infty} \cos\sqrt{\lambda}\, t\, d_\lambda f_s(y, \lambda) = \int_0^{\infty} \lambda \cos\sqrt{\lambda}\, t\, d_\lambda f_r(y, \lambda).$$

Again, suppose

$$\omega(\lambda, t, \Delta t) = \frac{1}{\lambda}\left[\frac{\cos\sqrt{\lambda}(t+\Delta t) - \cos\sqrt{\lambda}\,t}{\Delta t}\right]$$
$$= -\frac{1}{\sqrt{\lambda}}\sin\left[\sqrt{\lambda}\,t + \frac{\sqrt{\lambda}\,\Delta t}{2}\right]\frac{\sin\frac{1}{2}\sqrt{\lambda}\,\Delta t}{\frac{1}{2}\sqrt{\lambda}\,\Delta t}.$$

Then

$$(6.3)\quad \sum_{s=1}^{\infty} k_{rs}\int_0^{\lambda'} \omega(\lambda, t, \Delta t)\, d_\lambda f_s(y, \lambda)$$
$$= \int_0^{\lambda'} \frac{\cos\sqrt{\lambda}(t+\Delta t) - \cos\sqrt{\lambda}\,t}{\Delta t}\, d_\lambda f_r(y, \lambda).$$

From the inequality (6.2) its results that

$$\left|\sum_s k_{rs}\int_{\lambda'}^{\lambda''} \omega(\lambda, t, \Delta t)\, d_\lambda f_s^{(n)}(y, \lambda)\right| \leqq \frac{1}{\sqrt{\lambda'}}\sqrt{(y, y)\sum_s k_{rs}^2}$$

hence

$$\left|\sum_s k_{rs}\int_{\lambda'}^{\lambda''} \omega(\lambda, t, \Delta t)\, d_\lambda f_s(y, \lambda)\right|$$
$$= \left|\int_{\lambda'}^{\lambda''} \frac{1}{\Delta t}\left[\cos\sqrt{\lambda}(t+\Delta t) - \cos\sqrt{\lambda}\,t\right] d_\lambda f_r(y, \lambda)\right|$$
$$\leqq \frac{1}{\sqrt{\lambda'}}\sqrt{(y, y)\sum_s k_{rs}^2}.$$

Hence the integral on the right of (6.3) approaches its limit uniformly with respect to Δt as $\lambda' \to \infty$; the same is true of the sum on the left, and from Theorem II,

$$\lim_{\Delta t=0} \sum_{s=1}^{\infty} k_{rs}\int_0^{\lambda'} \omega(\lambda, t, \Delta t)\, d_\lambda f_s(y, \lambda) = -\sum_{s=1}^{\infty} k_{rs}\int_0^{\lambda'} \frac{1}{\sqrt{\lambda}}\sin\sqrt{\lambda}\,t\, d_\lambda f_s(y, \lambda)$$
$$= -\int_0^{\lambda'} \sqrt{\lambda}\sin\sqrt{\lambda}\,t\, d_\lambda f_r(y, \lambda).$$

Applying Theorem II again,

$$\sum_{s=1}^{\infty} k_{rs}\int_0^{\infty} \frac{\sin\sqrt{\lambda}\,t}{\sqrt{\lambda}}\, d_\lambda f_s(y, \lambda) = \int_0^{\infty} \sqrt{\lambda}\sin\sqrt{\lambda}\,t\, d_\lambda f_r(y, \lambda).$$

Suppose

$$x_s(t) = \int_0^{\infty} \cos\sqrt{\lambda}\,t\, d_\lambda f_s(y, \lambda);$$

from the uniform convergence it follows that

$$\frac{dx_s}{dt} = \lim_{\Delta t=0}\int_0^{\infty} \frac{\cos\sqrt{\lambda}(t+\Delta t) - \cos\sqrt{\lambda}\,t}{\Delta t}\, d_\lambda f_s(y, \lambda)$$
$$= -\int_0^{\infty} \sqrt{\lambda}\sin\sqrt{\lambda}\,t\, d_\lambda f_s(y, \lambda).$$

Suppose

$$\omega'(\lambda, t, \Delta t) = \frac{1}{\sqrt{\lambda}}\left[\frac{\sin\sqrt{\lambda}(t+\Delta t) - \sin\sqrt{\lambda}\,t}{\Delta t}\right]$$
$$= \cos\left(\sqrt{\lambda}\,t + \frac{\sqrt{\lambda}\,\Delta t}{2}\right)\frac{\sin\frac{1}{2}\sqrt{\lambda}\,\Delta t}{\frac{1}{2}\sqrt{\lambda}\,\Delta t}.$$

Then

$$\frac{x_s'(t+\Delta t) - x_s'(t)}{\Delta t} = -\sum_{i=1}^{\infty} k_{si}\int_0^{\infty}\omega'(\lambda, t, \Delta t)\, d_\lambda f_i(y, \lambda).$$

Now

$$\left|\int_{\lambda'}^{\lambda''}\omega'(\lambda, t, \Delta t)\, d_\lambda f_i(y, \lambda)\right| \leqq \text{Max}\,|\omega'|\int_{\lambda'}^{\lambda''}|df_i| \leqq \varepsilon$$

if λ' is such that $\int_{\lambda'}^{\infty}|df_i| < \varepsilon$ a condition which can be satisfied since the total variation of f_i is finite. Consequently the integral $\int_0^{\infty}\omega'(\lambda, t, \Delta t)\, d_\lambda f_i(y, \lambda)$ converges uniformly with respect to Δt, and

$$\lim_{\Delta t = 0}\int_0^{\infty}\omega'(\lambda, t, \Delta t)\, d_\lambda f_i(y, \lambda) = \int_0^{\infty}\cos\sqrt{\lambda}\,t\, d_\lambda f_i(y, \lambda).$$

$$\lim_{\Delta t = 0}\frac{x_s'(t+\Delta t) - x_s'(t)}{\Delta t} = -\sum_{i=1}^{\infty} k_{si}\int_0^{\infty}\cos\sqrt{\lambda}\,t\, d_\lambda f_i(y, \lambda),$$

$$\frac{d^2 x_s}{dt^2} = -\sum_{i=1}^{\infty} k_{si}\, x_i, \qquad s = 1, 2, \cdots, \infty.$$

Consider also the functions

$$y_r(t) = \int_0^{\infty}\frac{\sin\sqrt{\lambda}\,t}{\sqrt{\lambda}}\, d_\lambda f_r(\overline{y}, \lambda)$$

where $\overline{y}_1, \overline{y}_2, \cdots$ satisfy the condition $\sum_s \overline{y}_s^2$ converges; the integral

$$\int_0^{\infty}\frac{\sin\sqrt{\lambda}(t+\Delta t) - \sin\sqrt{\lambda}\,t}{\sqrt{\lambda}\,\Delta t}\, d_\lambda f_r(\overline{y}, \lambda)$$

converges uniformly with respect to Δt, hence

$$\frac{dy_r}{dt} = \int_0^{\infty}\cos\sqrt{\lambda}\,t\, d_\lambda f_r(\overline{y}, \lambda).$$

It has already been seen that this integral can be differentiated; hence

$$\frac{d^2 y_r}{dt^2} = -\int_0^{\infty}\sqrt{\lambda}\sin\sqrt{\lambda}\,t\, d_\lambda f_r(\overline{y}, \lambda) = -\sum_{s=1}^{\infty} k_{rs}\, y_s(t).$$

It will be shown that the conditions (6.11) are satisfied by the system of functions

$$z_r(t) = x_r(t, y') + y_r(t, y'').$$

If ε is given, l can be so chosen that

$$\left|\int_{\lambda'}^{\infty} \cos\sqrt{\lambda}\, t\, d_\lambda f_r(y', \lambda)\right| < \varepsilon, \qquad \lambda' > l.$$

Also,

$$\lim_{t=0}\int_0^{\lambda'} \cos\sqrt{\lambda}\, t\, d_\lambda f_r(y', \lambda) = \int_0^{\lambda'} d_\lambda f_r(y', \lambda) = y_r' - \int_{\lambda'}^{\infty} d_\lambda f_r(y', \lambda).$$

From the inequality above it is seen that the infinite integral

$$\int_0^{\infty} \cos\sqrt{\lambda}\, t\, d_\lambda f_r(y', \lambda)$$

converges uniformly with respect to t, for t real; if $\lambda' > l$,

$$\left|\int_{\lambda'}^{\infty} d_\lambda f_r(y', \lambda)\right| < \varepsilon,$$

consequently $\lim_{t=0} x_r(t, y') = y_r'$.

For a similar reason

$$\lim_{t=0}\int_0^{\infty} \frac{\sin\sqrt{\lambda}\, t}{\sqrt{\lambda}}\, d_\lambda f_r(y'', \lambda) = 0,$$

hence $\lim_{t=0} z_r(t) = y_r'$. Also

$$\frac{dx_r}{dt} = -\int_0^{\infty} \sqrt{\lambda}\sin\sqrt{\lambda}\, t\, d_\lambda f_r(y', \lambda) = -\sum_{s=1}^{\infty} k_{rs}\int_0^{\infty} \frac{\sin\sqrt{\lambda}\, t}{\sqrt{\lambda}}\, d_\lambda f_s(y', \lambda)$$

and from Theorem II,

$$\lim_{t=0}\frac{dx_r}{dt} = -\sum_{s=1}^{\infty} k_{rs}\lim_{t=0}\int_0^{\infty} \frac{\sin\sqrt{\lambda}\, t}{\sqrt{\lambda}}\, d_\lambda f_s(y', \lambda) = 0.$$

Consequently,

$$\lim_{t=0}\frac{dz_r}{dt} = \lim_{t=0}\frac{dy_r}{dt} = \lim_{t=0}\int_0^{\infty} \cos\sqrt{\lambda}\, t\, d_\lambda f_r(y'', \lambda) = y_r''.$$

It remains to prove that the series $\sum_r z_r^2$ converges. From the inequality $(x_r + y_r)^2 \leqq 2(x_r^2 + y_r^2)$ it is only necessary to show that the series $\sum_r x_r^2$, $\sum_r y_r^2$ converge. From the inequality (6.2),

$$\sum_{s=1}^{m} z_s^2 = \lim_{n\to\infty}\sum_{s=1}^{m} z_s^{(n)2} \leqq (y, y)\,\mathrm{Max}\,\omega(\lambda)^2, \qquad \lambda' \leqq \lambda \leqq \lambda''.$$

Consequently if $|\omega(\lambda)| \leqq M$, $0 \leqq \lambda < \infty$,

$$\sum_{s=1}^{m} \left(\int_0^\infty \omega(\lambda)\, d_\lambda f_s(y, \lambda) \right)^2 \leqq M^2 (y, y),$$

hence the series $\sum_{s=1}^{\infty} \left(\int_0^\infty \omega(\lambda)\, d_\lambda f_s(y, \lambda) \right)^2$ converges. Substituting $\omega(\lambda) = \cos \sqrt{\lambda}\, t$, with y_i replaced by y_i', one obtains the convergence of the series $\sum_s x_s^2$; substituting $\omega(\lambda) = (\sin \sqrt{\lambda}\, t)/\sqrt{\lambda}$ with y_i replaced by y_i'', one obtains the convergence of the series $\sum_s y_s^2$.

The results of this section may be expressed in the following theorem. *If k_{rs} $(r, s = 1, 2, \cdots, \infty)$ are real numbers which satisfy the conditions $k_{rs} = k_{sr}$, $\sum_s k_{rs}^2$ converges for every r, and if y_i' $(i = 1, \cdots, \infty)$, y_i'' $(i = 1, \cdots, \infty)$ are real numbers such that $\sum_s y_s'^2$, $\sum_s y_s''^2$ converge, then the set of equations and auxiliary conditions*

$$\frac{d^2 x_r}{dt^2} = -\sum_{s=1}^{\infty} k_{rs}\, x_s, \qquad r = 1, \cdots, \infty,$$

$$[x_r]_0 = y_r', \qquad \left[\frac{dx_r}{dt}\right]_0 = y_r'',$$

$$\sum_{s=1}^{\infty} x_r^2 \text{ converges},$$

possesses a solution, provided there exists a positive constant q which is less than any of the roots of the equations

$$\begin{vmatrix} k_{11} - \lambda & k_{12} & \cdots & k_{1n} \\ k_{21} & k_{22} - \lambda & \cdots & k_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ k_{n1} & k_{n2} & \cdots & k_{nn} - \lambda \end{vmatrix} = 0, \qquad n = 1, 2, \cdots, \infty.$$

One solution is represented by the sum

$$x_r = \int_0^\infty \cos \sqrt{\lambda}\, t\, d_\lambda f_r(y', \lambda) + \int_0^\infty \frac{\sin \sqrt{\lambda}\, t}{\sqrt{\lambda}}\, d_\lambda f_r(y'', \lambda),$$

in which $f_r(y', \lambda)$, $f_r(y'', \lambda)$ are constructed as in section 3.

The last condition stated in the theorem is equivalent to the condition that a positive constant p exist, for which the quadratic forms

$$\sum_{r,s=1}^{n} k_{rs}\, x_r\, x_s - p \sum_{s=1}^{n} x_s^2 = K_n(x) - p(x, x)_n$$

are positive definite for every n.

While the solution given above is determined by the construction of the functions $h_{rs}(\lambda)$, the preceding discussion gives no information concerning the uniqueness of the solution.

7. Applications to the theory of ordinary differential equations. The developments of the preceding sections can be applied to obtain formulae for the representation of the solution of a system of differential equations, in terms of a set of normalized, orthogonal functions chosen arbitrarily subject to certain restrictions.

Let $Q(u_1, \cdots, u_m, u_1', \cdots, u_m', x)$ be a quadratic form in the variables $u_i, u_i' = du_i/dx$, with coefficients which are continuous, with their derivatives of the first order, in the interval $0 \leqq x \leqq 1$. Let the set of functions $(u_1^{(r)}, u_2^{(r)}, \cdots, u_m^{(r)})$, $r = 1, 2, \cdots, \infty$ satisfy the conditions

$$u_i^{(r)}(0) = u_i^{(r)}(1) = 0, \qquad u_i^{(r)}(x), \quad \frac{du_i^{(r)}}{dx}, \quad \frac{d^2 u_i^{(r)}}{dx^2} \text{ are continuous,}$$

and

$$\int_0^1 \left\{ \sum_{i=1}^{m} u_i^{(r)} u_i^{(s)} \right\} dx = \delta_s^r.$$

A set of functions satisfying these conditions can be constructed from any normalized, orthogonal set $(u^{(r)})$ whose functions vanish at 0 and 1, and whose first two derivatives are continuous, $0 \leqq x \leqq 1$, by the choice

$$\begin{array}{llllllll}
u_1^{(1)} = u^{(1)}, & u_1^{(2)} = 0, & \cdots, & u_1^{(m)} = 0, & u_1^{(m+1)} = u^{(2)}, & \cdots, & u_1^{(2m)} = 0, & \cdots, \\
u_2^{(1)} = 0, & u_2^{(2)} = u^{(1)}, & \cdots, & u_2^{(m)} = 0, & u_2^{(m+1)} = 0, & & \vdots & \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \\
u_m^{(1)} = 0, & u_m^{(2)} = 0, & \cdots, & u_m^{(m)} = u^{(1)}, & u_m^{(m+1)} = 0, & \cdots, & u_m^{(2m)} = u^{(2)}, & \cdots.
\end{array}$$

The set of orthogonal functions will also be assumed to be complete, such that the only set of continuous functions $f_1(x), f_2(x), \cdots, f_m(x)$ which satisfy the conditions

$$\int_0^1 \left[\sum_{i=1}^{m} f_i(x) u_i^{(r)}(x) \right] dx = 0$$

for every r, is the set $f_i(x) = 0$, $i = 1, \cdots, m$.

Now suppose $f_i(x)$, $i = 1, \cdots, m$, a set of functions which are continuous for $0 \leqq x \leqq 1$; it is proposed to discuss the formal solution of the system

$$L_r(u) = \frac{d}{dx}\left(\frac{\partial Q}{\partial u_r'}\right) - \frac{\partial Q}{\partial u_r} = f_r(x), \qquad r = 1, \cdots, m, \tag{7.1}$$

such that

$$\int_0^1 Q(u, u', x)\,dx \geqq l \int_0^1 (u_1^2 + u_2^2 + \cdots + u_m^2)\,dx, \quad l > 0, \quad 0 \leqq x \leqq 1$$

for every set of functions $u_1(x), u_2(x), \cdots, u_n(x)$ which are continuous, together with their first order derivatives, on the interval $0 \leqq x \leqq 1$.

If $(u_1, u_2, \cdots, u_m)$, $(v_1, v_2, \cdots, v_m)$ are two sets of functions, of which each function is continuous, with its first two derivatives, on the given interval, and vanishes at the endpoints, and if $Q = Q(u)$, $Q' = Q(v)$, assume

$$Q(u, v) = \frac{1}{2}\sum_{i=1}^{m}\left[u_i \frac{\partial Q'}{\partial v_i} + u_i' \frac{\partial Q'}{\partial v_i'}\right] = \frac{1}{2}\sum_{i=1}^{m}\left[v_i \frac{\partial Q}{\partial u_i} + v_i' \frac{\partial Q}{\partial u_i'}\right].$$

Then

$$\begin{aligned}\int_0^1 Q(u, v)\,dx &= \frac{1}{2}\int_0^1\left[\sum_{i=1}^{m}\left(u_i \frac{\partial Q'}{\partial v_i} + u_i' \frac{\partial Q'}{\partial v_i'}\right)\right]dx \\ &= \frac{1}{2}\int_0^1\left[\sum_{i=1}^{m} u_i\left(\frac{\partial Q'}{\partial v_i} - \frac{d}{dx}\frac{\partial Q'}{\partial v_i'}\right)\right]dx + \frac{1}{2}\left[\sum_{i=1}^{m} u_i \frac{\partial Q'}{\partial v_i'}\right]_0^1 \\ &= -\int_0^1\left[\sum_{i=1}^{m} u_i L_i(v)\right]dx = -\int_0^1\left[\sum_{i=1}^{m} v_i L_i(u)\right]dx.\end{aligned}$$

Making the substitution

$$u_r(x) = \sum_{s=1}^{n} c_s u_r^{(s)}(x), \qquad r = 1, \cdots, m \tag{7.2}$$

one obtains

$$\int_0^1 Q(u)\,dx = -\int_0^1\left[\sum_{i=1}^{m} u_i L_i(u)\right]dx = +\sum_{r,s=1}^{n} c_r c_s \int_0^1 Q(u^{(r)}, u^{(s)})\,dx.$$

Suppose

$$k_{rs} = \int_0^1 Q(u^{(r)}, u^{(s)})\,dx = k_{sr},$$

then

$$\int_0^1 Q(u)\,dx = \sum_{r,s=1}^{n} k_{rs} c_r c_s.$$

Now if $\varphi_1(x), \cdots, \varphi_m(x)$ are a set of real functions which are continuous on the interval $0 \leqq x \leqq 1$, let

$$a_r = \int_0^1\left[\sum_{i=1}^{m} \varphi_i u_i^{(r)}\right]dx.$$

Then it is easily shown that

$$\sum_{r=1}^{\infty} a_r^2 \leqq \int_0^1\left[\sum_{i=1}^{m} \varphi_i^2\right]dx.$$

For,

$$\int_0^1 \left\{ \sum_{i=1}^{m} \left[\varphi_i - \sum_{r=1}^{n} a_r u_i^{(r)} \right]^2 \right\} dx \geqq 0,$$

from which

$$\int_0^1 \left[\sum_{i=1}^{m} \varphi_i(x)^2 \right] dx + 2 \sum_{\substack{k,r=1 \\ k \neq r}}^{n} a_k a_r \int_0^1 \left[\sum_{i=1}^{m} u_i^{(k)} u_i^{(r)} \right] dx$$
$$- 2 \sum_{r=1}^{n} a_r \int_0^1 \left[\sum_{i=1}^{m} \varphi_i u_i^{(r)} \right] dx$$
$$+ \sum_{r=1}^{n} a_r^2 \int_0^1 \left[\sum_{i=1}^{m} u_i^{(r)2} \right] dx \geqq 0$$

and when the integrals are evaluated the result stated follows. Since

$$k_{rs} = -\int_0^1 \left[\sum_i u_i^{(s)} L_i(u^{(r)}) \right] dx = k_{sr},$$

it follows that

$$\sum_{s=1}^{\infty} k_{rs}^2 \leqq \int_0^1 \left[\sum_{i=1}^{m} L_i(u^{(r)})^2 \right] dx$$

and similarly if

$$d_r = \int_0^1 \left[\sum_{i=1}^{m} u_i^{(r)} f_i(x) \right] dx, \qquad \sum_{s=1}^{\infty} d_s^2 \leqq \int_0^1 \left[\sum_{i=1}^{m} f_i(x)^2 \right] dx.$$

By hypothesis,

$$\int_0^1 Q(u)\, dx \geqq l \int_0^1 \left[\sum_{i=1}^{m} u_i^2 \right] dx$$

consequently

$$\sum_{r,s=1}^{n} k_{rs} c_r c_s \geqq l \sum_{s=1}^{n} c_s^2.$$

It follows that the roots of the equation $|k_{rs} - \delta_s^r \lambda|_n = 0$ are greater than the positive constant $l - \varepsilon$, for every n.

To obtain a formal solution of order n, of equations (7.1), let the substitution (7.2) be made; if the ith equation is multiplied by $u_i^{(r)}(x)$, and the sum formed,

$$\sum_{i=1}^{m} \int_0^1 u_i^{(r)} \left[\sum_{s=1}^{n} c_s L_i(u^{(s)}) \right] dx = d_r, \qquad r = 1, \cdots, n,$$

or

$$\sum_{s=1}^{n} k_{rs} c_s = -d_r.$$

If the order n is assumed to be infinite, the equations become

$$\sum_{s=1}^{\infty} k_{rs}\, c_s = -d_r, \qquad r = 1, 2, \cdots, \infty.$$

From the inequalities above, it is seen that this system is of the type discussed in section 5, with $\lambda = 0$; consequently a solution of the infinite system exists, and the series formed with the aid of this solution,

$$u_r(x) = \sum_{s=1}^{\infty} c_s\, u_s^{(r)}(x)$$

may be considered a formal solution of the equations (7.1).

8. **Application to a system of partial differential equations.** With the same definition of the linear expressions $L_i(u)$ as in the preceding section, except that $u_s' = \dfrac{\partial u_s}{\partial x}$, and the same hypotheses on the quadratic form Q, consider the system of partial differential equations and boundary conditions

$$\frac{\partial^2 u_r}{\partial t^2} = L_r(u), \qquad r = 1, \cdots, m, \tag{8.1}$$

$$u_r(0, t) = u_r(1, t) = 0.$$

In the first system of equations make the substitution

$$u_i(x, t) = \sum_{s=1}^{n} c_s(t)\, u_i^{(s)}(x), \qquad i = 1, \cdots, m.$$

It results immediately that

$$\sum_{i=1}^{m} \sum_{s=1}^{n} \frac{d^2 c_s}{dt^2} \int_0^1 u_i^{(s)}(x)\, u_i^{(r)}(x)\, dx = \sum_{s=1}^{n} c_s \int_0^1 \left[\sum_{i=1}^{m} u_i^{(r)}(x)\, L_i(u^{(s)}) \right] dx$$

from which

$$\frac{d^2 c_r}{dt^2} = -\sum_{s=1}^{n} k_{rs}\, c_s, \qquad r = 1, \cdots, n.$$

If the order n is assumed infinite,

$$\frac{d^2 c_r}{dt^2} = -\sum_{s=1}^{\infty} k_{rs}\, c_s, \qquad r = 1, \cdots, \infty.$$

From the inequalities of the preceding section it follows that the system satisfies the hypotheses of the theorem of section 6; a solution exists, and the corresponding system of series

$$u_r(x, t) = \sum_{s=1}^{\infty} c_s(t)\, u_r^{(s)}(x), \qquad r = 1, \cdots, m$$

may be considered a formal solution of the partial differential system (8.1).

THE GENERAL GEOMETRY OF PATHS.*

By Jesse Douglas.†

Introduction. The present paper develops the basis of the theory of the most general system of paths. By this term we denote any system of curves in an N-dimensional space S_N definable by equations of the form

$$x^i = f^i(t, a), \qquad i = 1, 2, \cdots, N, \tag{0.1}$$

where t is a parameter varying along each curve, and a denotes a set of $2N-2$ (essential) parameters $a_1, a_2, \cdots, a_{2N-2}$, which vary from curve to curve; where the functions f^i are analytic in all their arguments (analytic curves, analytically distributed); and where there is a unique curve of the system passing through any two given points not too far apart, and through any given point in any given direction.

Our principal result is that a system of differential equations of the form

$$\frac{d^2x^i}{dt^2} = \underset{2}{H^i}(x, p) \qquad \left(p = \frac{dx}{dt}\right) \tag{0.2}$$

where the functions $\underset{2}{H^i}$ are homogeneous of the second degree in p, is adequate to represent the most general system of paths.

In the work of J. L. Synge,‡ J. H. Taylor,§ and L. Berwald,‖ on general metric or Finsler¶ spaces—i. e., spaces where the element of length is defined to be $ds^2 = F(x, dx)$, F being any analytic function homogeneous of the second degree in the dx's—the geodesics of the space play an important part; their differential equations are a particular case of the form (0.2). But these writers in dealing with the extremals of the calculus of

* Received June 2, 1927; in revised form September 19, 1927.

† National Research Fellow in Mathematics.

‡ *A generalization of the Riemannian line-element,* Trans. Amer. Math. Soc., vol. 27 (1925), pp. 61–67.

§ *A generalization of Levi-Civita's parallelism and the Frenet formulas,* Trans. Amer. Math. Soc., vol. 27 (1925), pp. 246–264.

‖ (A) *Über Parallelübertragung in Räumen mit allgemeiner Massbestimmung,* Jahresb. Deutsch. Math.-Ver., vol. 34 (1925–26), pp. 213–220; (B) *Untersuchung über die Krümmung allgemeiner metrischer Räume* ···. Math. Zeitsch., vol. 25 (1926), pp. 40–73, (C) *Über zweidimensionale allgemeine metrische Räume,* Crelle, vol. 156 (1927), pp. 191–210.

¶ So called because the study of spaces of this type was begun by P. Finsler, *Über Kurven und Flächen in allgemeinen Räumen,* Dissertation, Göttingen, 1918.

variations problem $\int \sqrt{F(x, dx)} =$ minimum are considering only a special type of system of paths.

In the first section of Berwald's paper (A) the integral curves of a general system of differential equations of the form (0.2) occur, but the theorem that (0.2) has sufficient generality to represent any system of paths whatever appears for the first time in the present paper.

Our theory includes as a special case the theory of systems of curves definable by differential equations of the form

$$\frac{d^2x^i}{dt^2} = \Gamma^i_{jk}(x)\frac{dx^j}{dt}\frac{dx^k}{dt}, \tag{0.3}$$

i. e., where $H^i_2(x, p)$ are the simplest functions of their kind—homogeneous quadratic polynomials. This theory was initiated by L. P. Eisenhart and O. Veblen* and called by them the "geometry of paths". We shall reserve this term for the theory of the most general system of paths† as defined at the beginning of this paper, referring to the theory connected with (0.3) as the "restricted geometry of paths".

1. **Path, parameterized path.** It is vital in all our work to distinguish sharply between *curve* and *parameterized curve*. A curve is to be visualized as a thread, a parameterized curve as a thread together with a distribution of numbers over the points of the thread.

If we suppose S_N referred to a fixed coördinate system (x), then a parameterized curve is isomorphic with a set of equations

$$x^i = f^i(t), \tag{1.1}$$

where the f^i are analytic functions not all constants. A curve, on the other hand, is represented by

$$x^i = f^i[\varphi(t)], \tag{1.2}$$

where the f^i are fixed analytic functions not all constants, and φ ranges over all non-constant analytic functions. If φ is given a particular form

* See their notes in vol. 8 (1922) of Proc. Nat. Acad. Sci., also the notes of T. Y. Thomas and J. M. Thomas in vol. 11 (1925). A systematic account of this restricted geometry of paths is given in Veblen and T. Y. Thomas, *The geometry of paths*, Trans. Amer. Math. Soc., vol. 25 (1923), pp. 551–608.

All these writers take (0.3) in the form $\frac{d^2x^i}{dt^2} + \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0$—similarly, Berwald writes (0.2) as $\frac{d^2x^i}{dt^2} + H^i_2(x, p) = 0$. As a consequence, differences of sign appear in the formulas of the present paper as compared with those of the previous literature, but this need occasion no real confusion.

† Cf. Veblen, *Remarks on the foundations of geometry*, Bull. Amer. Math. Soc., vol. 31 (1925), p. 128.

in (1.2), a parameterized curve is defined, called a parameterization of the curve (1.2). A curve may be regarded as the class of all its parameterizations.

Similarly, we may distinguish between a system of parameterized paths and a system of paths. The equations (0.1) define a system of paths in a certain fixed parameterization. We may change this parameterization most generally by writing

$$t = \varphi(\tau, a), \tag{1.3}$$

the resulting equations

$$x^i = f^i[\varphi(\tau, a), a] \tag{1.4}$$

representing the same paths, each referred to a new parameter.

2. **The various geometries of paths—finite equations.** Imagine a system of paths realized as a system of threads in S_N. To represent this system of paths mathematically we use a reference system A consisting of the following three components:

(A_1) a coördinate system (x) for the containing space S_N,

(A_2) a parameter t for each path,

(A_3) a set of parameters $a = a_1, a_2, \cdots, a_{2N-2}$ to determine the various paths.

The three components of A being chosen, the system of paths is represented by a determinate set of equations

$$x^i = f^i(t, a) \tag{2.1}$$

obeying the conditions described at the beginning of this paper.

The system of paths remaining the same as a geometric configuration, the reference system A may be transformed most generally as follows:

(T_1) by a transformation of coördinates in S_N:

$$\bar{x}^i = \bar{x}^i(x),$$

(T_2) by a simultaneous transformation of parameter on all the paths:

$$t = \varphi(\tau, a),$$

(T_3) by a transformation of the parameters a:

$$a = a(\bar{a}).$$

Suppose (T_1), (T_2), (T_3) convert (2.1) into

$$\bar{x}^i = g^i(\tau, \bar{a}). \tag{2.2}$$

(2.2) is the representation in the new reference system B of the same system of paths.

The transformation (T_3) is of only minor importance since it does not affect the differential equations of the paths. All equations which express geometric properties of a system of paths must be invariant under an arbitrary transformation (T_3), and hereafter we take such invariance for granted.

We regard the paths as a collection of ∞^{2N-2} one-dimensional manifolds coexisting in S_N. Since the tripartite reference system A is something external to the system of paths, serving only to give it a definite mathematical representation, the geometry of the paths themselves, or, as we shall call it, the *descriptive* geometry of paths, is the theory of those properties of a set of equations (2.1) which are invariant under an arbitrary transformation (T_1) of coördinates, combined with an arbitrary transformation (T_2) of the parameterization of the paths, combined with an arbitrary transformation (T_3) of the a's.

This idea may be generalized by replacing the general groups of transformations (T_1), (T_2) by other less extensive groups. For instance, we may study the invariant theory of the equations (2.1) under the general group of coördinate transformations combined with the group

$$t = \alpha\tau + \beta \tag{2.3}$$

of simultaneous parameter transformations on the paths, where α, β are arbitrary functions of the a's. This theory may appropriately be termed the "affine geometry of paths", since, regarding the paths as so many one-dimensional manifolds each referred to the coördinate t, (2.3) represents an affine transformation of this coördinate. As will appear in § 4, this geometry is identical with that hitherto called (for the restricted case) the affine geometry of paths on account of its association with affinely connected manifolds.

Most generally expressed, the idea we are trying to bring out here is the following. Suppose that G is any group of coördinate transformations: affine, projective, descriptive,* etc. Further, let O denote an operation which consists of a simultaneous change of parameter on all the paths:

$$t = \varphi_0(\tau, a),$$

and let H denote any group of operations O. Then we may speak of the G, H geometry of paths: this is the invariant theory of the equations (2.1) under the transformations of G combined with those of H.

* By descriptive, we mean the group of all analytic transformations.

In the present paper G will always be the descriptive group. But geometries of paths where this is not the case have been considered: e. g., the geometry termed by T. Y. Thomas* the "equi-projective geometry of paths" (restricted case) is based on the group of coördinate transformations whose jacobian is unity, combined with the general group (T_2) of parameter transformations. In our nomenclature the name for this geometry would be the "equi-voluminar†-descriptive geometry of paths".

Another example of a geometry of paths is that based on the descriptive group (T_1) combined with

$$t = \tau + \beta \tag{2.4}$$

where β is an arbitrary function of the a's. The transformation (2.4) is evidently appropriate when the space which contains the system of paths is provided with a Riemann or Finsler metric, and the parameter t is the arc-length reckoned from an arbitrarily chosen point of each path. We shall see conversely (§ 3c) that any system of paths on which the parameter is determined up to the transformation (2.4) can be regarded as a system of paths in a Finsler space, the parameter being arc length. For this reason we shall refer to the geometry associated with (2.4) as the "metric geometry of paths".

We might cite further examples. For instance, the group of parameter transformations might be

$$t = \frac{\alpha\tau + \beta}{\gamma\tau + \delta} \tag{2.5}$$

where $\alpha, \beta, \gamma, \delta$ are arbitrary functions of the a's. The corresponding geometry is, in the present terminology, the "projective geometry of paths". In the literature on the restricted geometry of paths, the projective geometry of paths has meant the geometry of the paths themselves, apart from their generation by means of any particular affine connection, or what we have called the descriptive geometry of paths.

The present paper confines itself to the study of the following three geometries of paths: affine, descriptive, metric. Correspondingly, we shall use the terms affine, descriptive, metric space of paths for spaces of parameterized paths on which the parameter is subject respectively to the transformations (2.3), (T_2), (2.4).

---

* *On the projective and equi-projective geometries of paths*, Proc. Nat. Acad. Sci., vol. 11 (1925), pp. 199–204.

† Since, if we regard S_N for the moment as the bearer of a euclidean metric, transformations of jacobian unity are characterized by preserving volume.

3. The various geometries of paths—differential equations.

(a) *Affine.* Suppose given an affine space of paths

$$x^i = f^i(\alpha t+\beta, a). \tag{3a.1}$$

We proceed to find its differential equations.

Differentiating (3a.1) twice successively as to t, we get

$$p^i = \alpha f^{i'}(\alpha t+\beta, a), \tag{3a.2}$$

$$\frac{d^2x^i}{dt^2} = \alpha^2 f^{i''}(\alpha t+\beta, a). \tag{3a.3}$$

Since there is a unique path through each point in each direction, the $2n$ equations (3a.1), (3a.2) are solvable uniquely for the $2n$ unknowns α, $\alpha t+\beta$, a as functions of x, p. I say that on substituting these values in (3a.3) there will result a system of differential equations of the form

$$\frac{d^2x^i}{dt^2} = \underset{2}{H^i}(x, p), \tag{3a.4}$$

where the functions $\underset{2}{H^i}$ are homogenous of the second degree in p.

To prove this, observe that if in (3a.2) each p^i is replaced by λp^i, the resulting equations may be written

$$p^i = \frac{\alpha}{\lambda} f^{i'}(\alpha t+\beta, a); \tag{3a.2'}$$

and (3a.1), (3a.2') are the same equations in α/λ, $\alpha t+\beta$, a that (3a.1), (3a.2) are in α, $\alpha t+\beta$, a. Hence $\alpha t+\beta$, a have the same values as before, while α has λ times the value it had before. On account of the factor α^2 in (3a.3), it follows that the effect of multiplying each p by λ is to multiply d^2x^i/dt^2 by λ^2—but this is the definition of homogeneity of the second degree.

Conversely, suppose given any system of differential equations of the form (3a.4). Since (3a.4) is invariant under the transformation $t = \alpha\tau+\beta$, its general integral has the form (3a.1), i. e., (3a.4) represents an affine space of paths.

(b) *Descriptive.* The finite equations of a descriptive space of paths are

$$x^i = f^i[\varphi(t, a), a], \tag{3b.1}$$

where φ is an arbitrary function of its arguments.

Let us imbed each parameterization φ of the paths in an affine space of paths by writing $\alpha t+\beta$ in place of t:

$$x^i = f^i[\varphi(\alpha t+\beta, a), a]. \tag{3b.2}$$

Differentiating (3b.2) twice successively with respect to t, we get

$$p^i = \alpha\varphi' f^{i'}[\varphi(\alpha t+\beta, a), a], \tag{3b.3}$$

$$\frac{d^2x^i}{dt^2} = \alpha^2\varphi'^2 f^{i''}[\varphi(\alpha t+\beta, a), a]+\alpha^2\varphi'' f^{i'}[\varphi(\alpha t+\beta, a), a]. \tag{3b.4}$$

Observe that, if we omit for the moment the second term on the right of (3b.4), we have the same elimination theory as in (3a) with $\alpha\varphi'$, φ, a in place of α, $\alpha t+\beta$, a. It follows that after elimination the first term on the right of (3b.4) takes the form $\underset{2}{H^i}(x, p)$. Moreover the form of $\underset{2}{H^i}$ is independent of the form of φ as function of $\alpha t+\beta$, a, since $\alpha\varphi'$, φ figure as whole symbols in the process of elimination. The functions $\underset{2}{H^i}$ are determined by the arbitrarily chosen initial parameterization of the paths corresponding to $\varphi(t, a) = t$ in (3b.1).

The second term on the right of (3b.4) can be written

$$p^i \cdot \alpha\frac{\varphi''(\alpha t+\beta, a)}{\varphi'(\alpha t+\beta, a)}. \tag{3b.5}$$

Solving the $2n$ equations (3b.2), (3b.3) for the $2n$ unknowns α, $\alpha t+\beta$, a, and substituting in (3b.5), this takes the form

$$p^i G(x, p), \tag{3b.6}$$

where G is independent of the index i but depends on the form of φ, i. e., on the parameterization of the paths.

It is readily shown that G must be homogeneous of the first degree in p. For, as was seen in (3a), the effect of multiplying each p by λ is to multiply α by λ and leave $\alpha t+\beta$, a the same; this evidently has the effect of multiplying $G(x, p)$, equal to the second factor of (3b.5), by λ.

Summarizing, we have for the differential equations of a descriptive space of paths:

$$\frac{d^2x^i}{dt^2} = \underset{2}{H^i}(x, p)+p^i\underset{1}{H}(x, p), \tag{3b.7}$$

where the $\underset{2}{H^i}$ are *fixed* functions of their kind—determined by an arbitrarily chosen initial parameterization of the paths—and $\underset{1}{H}$ is an *arbitrary** function of its kind, varying with the parameterization of the paths.

* That $\underset{1}{H}$ is arbitrary of its kind follows from (3b.10) below, which gives the parameter corresponding to any assigned $\underset{1}{H}$.

Thus the two systems of differential equations

$$(3b.8)\quad \frac{d^2x^i}{dt^2} = \underset{2}{H^i}\left(x, \frac{dx}{dt}\right), \qquad \frac{d^2x^i}{d\tau^2} = \underset{2}{H^i}\left(x, \frac{dx}{d\tau}\right) + \frac{dx^i}{d\tau}\underset{1}{H}\left(x, \frac{dx}{d\tau}\right)$$

represent the same system of paths in two different parameterizations. If t, τ denote the corresponding parameters on an arbitrary path, we have by (3b.5)

$$(3b.9)\quad \frac{\alpha\varphi''}{\varphi'} = \frac{d^2t/d\tau^2}{dt/d\tau} = \underset{1}{H}\left(x, \frac{dx}{d\tau}\right),$$

whence

$$(3b.10)\quad \tau = \int e^{-\int \underset{1}{H}(x, dx)}\, dt$$

the integrals being taken over the arbitrary path.

(c) *Metric.* The finite equations of a metric space of paths are*

$$(3c.1)\quad x^i = f^i(t+\beta, a).$$

By two successive differentiations as to t, we get

$$(3c.2)\quad p^i = f^{i'}(t+\beta, a),$$

$$(3c.3)\quad \frac{d^2x^i}{dt^2} = f^{i''}(t+\beta, a).$$

The system (3c.1), (3c.2) consists of $2n$ equations in the $2n-1$ unknowns $t+\beta, a$. We can solve these equations for $t+\beta, a$ as functions of x, p, and obtain besides a relation in x, p:

$$(3c.4)\quad \theta(x, p) = 0.$$

The function θ must be non-homogeneous in p; otherwise the ratios of the p's would be restricted, contrary to our hypothesis of the existence of a path through each point in each direction. Further, there cannot be a second relation of the form (3c.4), independent of that one, for then again a restriction would result on the ratios of the p's.

On substituting in (3c.3) the values of $t+\beta, a$ as functions of x, p obtained from (3c.1), (3c.2), we get a system of differential equations of the form

$$(3c.5)\quad \frac{d^2x^i}{dt^2} = G^i(x, p).$$

It is possible—and in only one way—to normalize the system (3c.4), (3c.5) to the form

* In § 3c (see also 2·4) $t+\beta$ may be replaced without change of result by $\pm t+\beta$ (the arc length on a path may be reckoned in either sense).

$$F(x, p) = 1, \tag{3c.6}$$

$$\frac{d^2x^i}{dt^2} = \underset{2}{H^i}(x, p), \tag{3c.7}$$

where F and $\underset{2}{H^i}$ are homogeneous of the second degree in p. We mean by this that there exist uniquely determined functions F, $\underset{2}{H^i}$ such that

$$F(x, p) = 1, \qquad \underset{2}{H^i}(x, p) = G^i(x, p) \tag{3c.8}$$

for all values of x, p obeying $\theta(x, p) = 0$.

For the proof, suppose the x's to have arbitrary fixed values, and let the p's be imagined as rectangular coördinates in an auxiliary Cartesian space. Then $\theta(x, p) = 0$ represents a hypersurface Σ in this space, which on account of the non-homogeneity of θ, is not a cone with vertex at the origin O. The rays which join O to the various points P of Σ therefore cover a certain region of the p-space, containing Σ. Let us form a function $\underset{2}{H^i}(Q)$ of the points Q of this region by assigning to Q the value $\lambda^2 G^i(P)$, where $G^i(P)$ is the value of the function G^i for the point P of Σ which determines the ray OPQ, and $\lambda = OQ/OP$. Evidently $\underset{2}{H^i}$ is homogeneous of the second degree in p, and is equal to G^i where $\lambda = 1$ —i. e., on Σ, where $\theta(x, p) = 0$.

Similarly we can prove the unique existence of a function $F(x, p)$ homogeneous of the second degree in p and equal to unity where $\theta(x, p) = 0$.

Differentiating (3c.6) as to t and using (3c.7), we find

$$\frac{\partial F}{\partial x^i} p^i + \frac{\partial F}{\partial p^i} \underset{2}{H^i} = 0. \tag{3c.9}$$

This relation signifies that any one of the differential equations (3c.7) is a consequence of the other $n - 1$ and of (3c.6).

Conversely, given a system of differential equations of the form (3c.6), (3c.7) obeying the relation (3c.9), the general integral of these equations has the form (3c.1), since the system is invariant under the transformation $t = \tau + \beta$.

Suppose we define a Finsler metric in S_N by writing

$$ds^2 = F(x, dx). \tag{3c.10}$$

We have from (3c.6), $dt^2 = F(x, dx)$, so that $ds^2 = dt^2$, or $s = \pm t + \beta$; i. e., the parameter t is arc-length measured from an arbitrary initial point on each path in either sense.

Defining, with Synge,

$$g_{ij} = \frac{1}{2}\frac{\partial^2 F(x,p)}{\partial p^i \partial p^j}, \tag{3c.11}$$

we may, with help of Euler's theorem on homogeneous functions, write (3c.10) in a form like that of a Riemann metric:

$$ds^2 = g_{ij}\,dx^i\,dx^j, \tag{3c.12}$$

or (3c.6) in the form

$$g_{ij}\,p^i\,p^j = 1. \tag{3c.13}$$

The differential equations of the geodesics of (3c.12) are*

$$\frac{d^2x^i}{ds^2} = \left\{\begin{matrix} i \\ jk \end{matrix}\right\} p^j\,p^k, \tag{3c.14}$$

where

$$\left\{\begin{matrix} i \\ jk \end{matrix}\right\} = \frac{1}{2}\,g^{ia}\left(\frac{\partial g_{jk}}{\partial x^a} - \frac{\partial g_{ak}}{\partial x^j} - \frac{\partial g_{ja}}{\partial x^k}\right). \tag{3c.15}$$

It is convenient to write (3c.7) in the form

$$\frac{d^2x^i}{ds^2} = \left\{\begin{matrix} i \\ jk \end{matrix}\right\} p^j\,p^k + C^i(x,p) \tag{3c.16}$$

$-C^i$ is then the curvature vector of the path through the point x in the direction p. The relation (3c.9) becomes

$$g_{ij}\,p^i\,C^j = 0, \tag{3c.17}$$

expressing the transversality of the curvature vector to the direction p, or (Synge, § 5) the perpendicularity of C with respect to p.

(d) *The various geometries as invariant theories.* On differentiating twice with respect to t the equations

$$\bar{x}^i = \bar{x}^i(x), \tag{3d.1}$$

expressing an arbitrary transformation of coördinates, we get

$$\bar{p}^i = \frac{\partial \bar{x}^i}{\partial x^j}\,p^j, \tag{3d.2}$$

$$\underset{2}{\bar{H}}{}^i = \frac{\partial \bar{x}^i}{\partial x^j}\,\underset{2}{H}{}^j + \frac{\partial^2 \bar{x}^i}{\partial x^j\,\partial x^k}\,p^j\,p^k. \tag{3d.3}$$

* Synge, loc. cit., § 3.

Thus the affine geometry of paths is, analytically, the invariant theory of a set of functions $\underset{2}{H^i}(x, p)$ under the group of transformations (3d.1, 2, 3).

The descriptive geometry of paths is the invariant theory of a set of functions $\underset{2}{H^i}(x, p)$ under the group obtained by adjoining to (3d.1, 2, 3) the transformation

$$'\underset{2}{H^i} = \underset{2}{H^i} + p^i \underset{1}{H}, \tag{3d.4}$$

expressive of an arbitrary change of parameter on the paths.

The metric geometry of paths is the invariant theory under arbitrary transformation of coördinates of (1) a tensor $g_{ij}(x, p)$, homogeneous of degree zero in p, symmetric in i, j, and such that $g_{ijk} = \partial g_{ij}/\partial p^k$ is symmetric in i, j, k, together with (2) a vector $C^i(x, p)$ homogeneous of the second degree in p, the tensor and vector being connected by the relation $g_{ij} p^i C^j = 0$.

4. **Affinely connected manifolds.** In this section we give a second justification for the name "affine geometry of paths", showing how what we have called in § 2 an affine space of paths is related to the idea of affinely connected manifold in the sense defined for the restricted theory by Weyl.

In the most general sense, a space S_N is affinely connected if, given any two points A and B together with a curve AB, there is determined a linear correspondence between contravariant vectors ξ_A at A and contravariant vectors ξ_B at B; i. e.

$$\xi_B = \varphi(\xi_A, AB), \tag{4.1}$$

the function φ being such that for arbitrary scalar multipliers λ, λ':

$$\varphi(\lambda \xi_A + \lambda' \xi'_A, AB) = \lambda \varphi(\xi_A, AB) + \lambda' \varphi(\xi'_A, AB). \tag{4.2}$$

The vector ξ_B is said to be derived from ξ_A by parallel displacement along the curve AB.

Let us determine all affine connections definable in the following way by differential equations of the first order. Write

$$\frac{d\xi^i}{dt} = F^i\left(x, \frac{dx}{dt}, \xi\right); \tag{4.3}$$

then on the parameterized curve AB the x^i will be known functions of t, $x^i = x^i(t)$, which when substituted in (4.3) will reduce it to a system of differential equations for the ξ^i as functions of t. Suppose that A corresponds to $t = t_0$, B to a fixed general value of t; then solve the equations (4.3)

under the initial conditions: for $t = t_0$, $\xi^i = \xi^i_A$. This will determine the ξ^i as functions of t, and thus fix at any point B of the curve $x^i = x^i(t)$ a vector ξ_B, which we define as the one into which ξ_A is converted by our parallel displacement.

If the linear law (4.2) is to obtain, the differential equations (4.3) must be linear in ξ:

$$\frac{d\xi^i}{dt} = G^i_j\left(x, \frac{dx}{dt}\right)\xi^j. \tag{4.4}$$

A natural further requirement is that the affine connection be independent of the parameter used on the curve AB. If we replace t by $\varphi(t)$ in (4.4), this becomes

$$\frac{d\xi^i}{dt} = \varphi'(t)\, G^i_j\left(x, \frac{1}{\varphi'(t)}\frac{dx}{dt}\right)\xi^j; \tag{4.5}$$

and the necessary and sufficient condition that (4.5) be independent of $\varphi(t)$ is that the functions G^i_j be homogeneous of the first degree in the arguments dx/dt:

$$\frac{d\xi^i}{dt} = \underset{1}{H}{}^i_j\left(x, \frac{dx}{dt}\right)\xi^j, \tag{4.6}$$

or, using differentials instead of derivatives, as the homogeneity of the first degree permits us to do,

$$d\xi^i = \underset{1}{H}{}^i_j(x, dx)\,\xi^j. \tag{4.7}$$

Let us define

$$\Gamma^i_{jk}(x, dx) = \frac{\partial \underset{1}{H}{}^i_j(x, dx)}{\partial\, dx^k}; \tag{4.8}$$

then by Euler's theorem (4.7) can be written

$$d\xi^i = \Gamma^i_{jk}(x, dx)\, dx^k\, \xi^j. \tag{4.9}$$

The Γ^i_{jk} are homogeneous of degree zero in dx, i. e., they depend only on the ratios of the dx's (functions of position-direction).

In particular, if the Γ^i_{jk} are functions of the x's alone, then (4.9) defines the Weyl type of affine connection. By analogy with this special case we term $\Gamma^i_{jk}(x, dx)$ the *components of affine connection* in the coördinate system (x).

Let us specialize our affine connection (4.9) further by requiring its components Γ^i_{jk} to be symmetric in j, k, i. e., by (4.8),

$$\frac{\partial \underset{1}{H}{}^i_j}{\partial\, dx^k} = \frac{\partial \underset{1}{H}{}^i_k}{\partial\, dx^j}. \tag{4.10}$$

These are the necessary and sufficient conditions for the existence for each i of a function $\underset{2}{H^i}(x, dx)$ homogeneous of the second degree in dx, such that

$$\underset{1}{H^i_j} = \frac{1}{2}\frac{\partial \underset{2}{H^i}}{\partial dx^j}. \tag{4.11}$$

Then, by (4.8),

$$\Gamma^i_{jk} = \frac{1}{2}\frac{\partial^2 \underset{2}{H^i}}{\partial dx^j \partial dx^k}; \tag{4.12}$$

so that we have finally for the most general symmetric affine connection definable as described by differential equations of the first order:

$$d\xi^i = \frac{1}{2}\frac{\partial^2 \underset{2}{H^i}}{\partial dx^j \partial dx^k} dx^j \xi^k. \tag{4.13}$$

Let us inquire as to the parameterized curves whose tangent vectors dx/dt are parallel under the affine connection (4.13): they will be called the *parameterized geodesics* of the affine connection. Using derivatives instead of differentials in (4.13), writing $\xi = dx/dt$, and using Euler's theorem, we find for these geodesics:

$$\frac{d^2x^i}{dt^2} = \underset{2}{H^i}\left(x, \frac{dx}{dt}\right). \tag{4.14}$$

That is, we have the theorem: *the parameterized geodesics of the most general symmetric affine connection definable by differential equations of the first order form an affine space of paths in the sense of* (3a).

Conversely, if we start from any affine space of paths (4.14), there is determined by (4.13) a unique symmetric affine connection of which the given paths are the parameterized geodesics.

The law of transformation of the Γ^i_{jk} under change of the coördinate system is to be found by applying the operator $\frac{1}{2}\partial^2/\partial p^j \partial p^k$ to (3d.3); this gives

$$\frac{\partial \bar{x}^a}{\partial x^j}\frac{\partial \bar{x}^b}{\partial x^k}\bar{\Gamma}^i_{ab} = \frac{\partial \bar{x}^i}{\partial x^a}\Gamma^a_{jk} + \frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k}, \tag{4.15}$$

or solved for the $\bar{\Gamma}$'s,

$$\bar{\Gamma}^i_{jk} = \frac{\partial x^a}{\partial \bar{x}^j}\frac{\partial x^b}{\partial \bar{x}^k}\left(\frac{\partial \bar{x}^i}{\partial x^c}\Gamma^c_{ab} + \frac{\partial^2 \bar{x}^i}{\partial x^a \partial x^b}\right), \tag{4.16}$$

which is identical with the well-known law for the restricted case $\Gamma^i_{jk} = \Gamma^i_{jk}(x)$.

This is the convenient place to introduce the tensor

$$H^i_{jkl} = \frac{\partial \Gamma^i_{jk}}{\partial p^l} = \frac{\partial \Gamma^i_{jl}}{\partial p^k} = \frac{\partial \Gamma^i_{kl}}{\partial p^j} = \frac{1}{2}\frac{\partial^3 \underset{2}{H^i}}{\partial p^j \partial p^k \partial p^l}. \tag{4.17}$$

The tensor character of H^i_{jkl} is seen by operating on (4.15) with $\partial/\partial p^l$. Evidently the restricted affine geometry of paths is characterized by the condition

$$H^i_{jkl} = 0. \tag{4.18}$$

5. **The fundamental descriptive invariant.** An arbitrary change of parameter on the paths is expressed by the equation

$$'\underset{2}{H^i} = \underset{2}{H^i} + p^i \underset{1}{H}. \tag{5.1}$$

Applying to this the operator $\frac{1}{2}\partial^2/\partial p^j \partial p^k$, we find that the effect on the Γ's of a parameter transformation is

$$'\Gamma^i_{jk} = \Gamma^i_{jk} + \delta^i_j \underset{1}{H_k} + \delta^i_k \underset{1}{H_j} + p^i \underset{1}{H_{jk}}, \tag{5.2}$$

where

$$\underset{1}{H_j} = \frac{1}{2}\frac{\partial \underset{1}{H}}{\partial p^j}, \qquad \underset{1}{H_{jk}} = \frac{1}{2}\frac{\partial^2 \underset{1}{H}}{\partial p^j \partial p^k}. \tag{5.3}$$

With the help of Euler's theorem, (5.2) implies for the contracted Γ's:

$$'\Gamma^a_{ak} = \Gamma^a_{ak} + (N+1)\underset{1}{H_k}. \tag{5.4}$$

Differentiating this with respect to p^j, we get

$$'H^a_{jka} = H^a_{jka} + (N+1)\underset{1}{H_{jk}}. \tag{5.5}$$

It is easy to eliminate $\underset{1}{H}$ from (5.2), (5.4), (5.5); we find

$$\begin{aligned} '\Gamma^i_{jk} - \frac{\delta^i_j}{N+1}\,'\Gamma^a_{ak} - \frac{\delta^i_k}{N+1}\,'\Gamma^a_{ja} - \frac{1}{N+1}\,p^i\,'H^a_{jka} \\ = \Gamma^i_{jk} - \frac{\delta^i_j}{N+1}\Gamma^a_{ak} - \frac{\delta^i_k}{N+1}\Gamma^a_{ja} - \frac{1}{N+1}\,p^i H^a_{jka}. \end{aligned} \tag{5.6}$$

In other words, the quantities

$$\Pi^i_{jk} = \Gamma^i_{jk} - \frac{\delta^i_j}{N+1}\Gamma^a_{ak} - \frac{\delta^i_k}{N+1}\Gamma^a_{ja} - \frac{1}{N+1}\,p^i\,H^a_{jka} \tag{5.7}$$

depend only on the paths themselves, regardless of their parameterization. We term the invariant* which consists in the association with each coördinate system (x) of the set of functions $\Pi^i_{jk}(x, p)$ the *fundamental descriptive invariant*, because, as we shall see (§ 7), every differential invariant which depends only on the paths themselves is expressible in terms of Π^i_{jk} and its partial derivatives.

The invariant Π^i_{jk} was introduced into the restricted theory by T. Y. Thomas (loc. cit.) and termed by him the "projective connection". In the restricted theory the last term vanishes from the expression for Π^i_{jk}.

The law of transformation of the Π's under transformation of the coördinate system is identical with that of the restricted theory. For, by contraction of (4.16) we get

$$\bar{\Gamma}^a_{ak} = \frac{\partial x^b}{\partial \bar{x}^k}\left(\Gamma^a_{ab} + \frac{\partial \log \Delta}{\partial x^b}\right), \tag{5.8}$$

where Δ is the jacobian $|\partial \bar{x}/\partial x|$. Using the expression (5.7) for the Π's, and combining (4.15), (5.8), and the tensor law of transformation of H^a_{jka}, we find

$$\frac{\partial \bar{x}^a}{\partial x^j}\frac{\partial \bar{x}^b}{\partial x^k}\bar{\Pi}^i_{ab} = \frac{\partial \bar{x}^i}{\partial x^a}\Pi^a_{jk} + \frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k} - \frac{\partial \bar{x}^i}{\partial x^j}\frac{\partial \theta}{\partial x^k} - \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial \theta}{\partial x^j}, \tag{5.9}$$

where

$$\theta = \frac{1}{n+1}\log \Delta.$$

Applying to (5.9) the operator $\partial/\partial p^l$, we get an equation which expresses that

$$\mathfrak{H}^i_{jkl} = \frac{\partial \Pi^i_{jk}}{\partial p^l} = \frac{\partial \Pi^i_{jl}}{\partial p^k} = \frac{\partial \Pi^i_{kl}}{\partial p^j} \tag{5.10}$$

are the components of a tensor. Let us say that a system of paths is descriptively of the restricted type if there exists a parameterization of the paths for which the H^i are quadratic polynomials, i. e., for which $H^i_{jkl} = 0$. Differentiating (5.7) with respect to p^l, we have

$$\mathfrak{H}^i_{jkl} = H^i_{jkl} - P\left(\frac{\delta^i_j}{N+1}H^a_{akl}\right) - \frac{1}{N+1}p^i H^a_{jkla}, \tag{5.11}$$

where P indicates the sum of the three terms obtained by cyclic permutation of j, k, l. Hence, if for some parameterization of the paths, $H^i_{jkl} = 0$, it follows that $\mathfrak{H}^i_{jkl} = 0$.

* For the general idea of invariant in the sense used here see § 7.

Conversely, suppose $\mathfrak{H}^i_{jkl} = 0$. Operating on (5.2) with $\partial/\partial p^l$, we find the effect on H^i_{jkl} of a parameter transformation to be

$$'H^i_{jkl} = H^i_{jkl} + P(\delta^i_j \underset{1}{H}_{kl}) + p^i \underset{1}{H}_{jkl}. \tag{5.12}$$

Choosing

$$\underset{1}{H} = -\frac{2}{N+1} \Gamma^a_{ak} p^k, \tag{5.13}$$

we find

$$'H^i_{jkl} = \mathfrak{H}^i_{jkl} = 0, \tag{5.14}$$

i. e., there exists a parameterization of the paths for which the tensor H^i_{jkl} vanishes.

6. **The canonical parameter.** Suppose we have a system of paths referred to a fixed coördinate system (x); then a system of functions $\Pi^i_{jk}(x,p)$ is determined. Also, for every parameterization assigned to the paths there is determined an affine connection $\Gamma^i_{jk}(x,p)$. We now prove that corresponding to each coördinate system (x), there exists a parameter t_x, determined up to linear transformation, such that the Γ's determined by this parameter are equal to the Π's. t_x will be called the *canonical parameter* relative to the coördinate system (x)—in the restricted theory t_x is identical with the "projective parameter" of T. Y. Thomas (loc. cit.).

For the proof, let us start with an arbitrary parameterization of the paths; this will determine an affine connection Γ^i_{jk}. Then let us make the change of parameter defined by (5.13): we get by (5.2)

$$'\Gamma^i_{jk} = \Gamma^i_{jk} - \frac{\delta^i_j}{N+1}\Gamma^a_{ak} - \frac{\delta^i_k}{N+1}\Gamma^a_{ja} - \frac{1}{N+1} p^i H^a_{jka} = \Pi^i_{jk}, \tag{6.1}$$

which was the condition to be satisfied.

Since, as can be verified from (5.7), $\Pi^a_{ak} = 0$, the contracted Γ's are zero for the canonical parameter. Furthermore, this property is characteristic of the canonical parameter, for if in (5.4) we have $\Gamma^a_{ak} = 0$, $'\Gamma^a_{ak} = 0$, it follows that $\underset{1}{H}_k = 0$, and therefore $\underset{1}{H} = 2 p^k \underset{1}{H}_k = 0$. This implies that the canonical parameter is determined up to linear transformation.

What is the effect on the canonical parameter of a transformation of the coördinate system? Combining (5.4) and (5.8), we find that the effect on the contracted Γ's of a change of parameter followed by a transformation of coördinates is

$$'\bar{\Gamma}^a_{ak} = \frac{\partial x^b}{\partial \bar{x}^k}\left(\Gamma^a_{ab} + (N+1)\underset{1}{H}_b + \frac{\partial \log \Delta}{\partial x^b}\right). \tag{6.2}$$

Hence if ${}'\bar{\Gamma}^a_{ak} = 0$ as well as $\Gamma^a_{ab} = 0$, we have

$$\underset{1}{H}_k = -\frac{1}{N+1}\frac{\partial \log \Delta}{\partial x^k};$$

therefore

$$\underset{1}{H} = 2p^k \underset{1}{H}_k = -\frac{2}{N+1}\frac{\partial \log \Delta}{\partial x^k}\, p^k. \tag{6.3}$$

It follows by (3b.10) that

$$t_{\bar{x}} = \int e^{\int \frac{2}{N+1}\frac{\partial \log \Delta}{\partial x^k}\, dx^k} \cdot dt_x = \int \Delta^{\frac{2}{N+1}}\, dt_x. \tag{6.4}$$

A consequence of (6.4) is that the canonical parameter remains the same for all and only those transformations for which $\Delta =$ constant. Thus a geometry of paths which makes the canonical parameter fundamental must be based on the group of coördinate transformations of constant jacobian.

7. **Differential invariants, tensors.** The idea of invariant has been defined by O. Veblen* as follows. Suppose we have a space S_N referred to different coördinate systems (x), $(\bar{x})$, $(\tilde{x}), \cdots$. Suppose further that with each coördinate system (x) there is associated a definite mathematical configuration R: e. g., a system of differential equations, a set of functions, etc. This association or correspondence is called an *invariant*—abstractly the invariant may be identified with the class of pairs $[(x), R]$, $[(\bar{x}), \bar{R}]$, $[(\tilde{x}), \tilde{R}], \cdots$. R is called the *representation* of the invariant in the coördinate system (x).†

The mathematical nature of the invariant depends on the law of transformation of R when the coördinate system is changed, i. e., the law which, given (x), R, and $(\bar{x})$, enables us to determine $\bar{R}$:

$$\bar{R} = \mathfrak{F}[R, (x), (\bar{x})], \tag{7.1}$$

where the functional symbol $\mathfrak{F}$ is to be understood in the most general sense. This law of transformation must have the following group property: suppose we deduce $\bar{R}$ from R by (7.1), and then $\tilde{R}$ from $\bar{R}$ by

$$R = \mathfrak{F}[\bar{R}, (\bar{x}), (\tilde{x})]; \tag{7.2}$$

the result must be the same as if we deduced $\tilde{R}$ directly from R:

$$\tilde{R} = \mathfrak{F}[R, (x), (\tilde{x})]. \tag{7.3}$$

* See his Cambridge Tract (1927), *Invariants of Quadratic Differential Forms.*

† The idea of invariant may be generalized by having instead of a coördinate system (x) more general reference systems, such as that of § 2.

Thus any invariant may be defined by associating with an initial coördinate system (x) some mathematical configuration R, and then assigning a law of transformation (7.1) having the requisite group property. For example, an affine connection may be defined by associating with some fixed coördinate system (x) a set of functions $\Gamma^i_{jk}(x, p)$ having the necessary and sufficient properties of being (1) homogeneous of degree zero in p, (2) symmetric in j, k, (3) such that $H^i_{jkl} = \partial \Gamma^i_{jk}/\partial p^l$ is symmetric in j, k, l; and then assigning the law of transformation (4.16). Or, the descriptive invariant Π may be defined by means of an initial set of functions $\Pi^i_{jk}(x, p)$, together with the law of transformation (5.9).

In connection with the geometry of paths we shall distinguish three types of invariant, to be termed respectively *affine, descriptive, metric, differential invariants.*

An affine differential invariant will mean a set of functions of x, p, the Γ's, and the partial derivatives of the Γ's up to any finite order:

$$F^{a_1 \cdots a_r}_{b_1 \cdots b_s}\left[x, p, \Gamma, \frac{\partial \Gamma}{\partial x}, \frac{\partial \Gamma}{\partial p}, \cdots\right], \tag{7.4}$$

which are related to the same functions formed for $\bar{x}, \bar{p}, \bar{\Gamma}$, and the derivatives of $\bar{\Gamma}$:

$$F^{a_1 \cdots a_r}_{b_1 \cdots b_s}\left[\bar{x}, \bar{p}, \bar{\Gamma}, \frac{\partial \bar{\Gamma}}{\partial \bar{x}}, \frac{\partial \bar{\Gamma}}{\partial \bar{p}}, \cdots\right], \tag{7.5}$$

by any prescribed law of transformation (7.1). In particular, if the law of transformation is the tensor law (see 9.1), then we have an affine tensor differential invariant, or simply an *affine tensor.* The coördinates x can not appear explicitly in the expression for an affine tensor, as can be seen by considering the transformation $\bar{x}^i = x^i + a^i$, where the a^i are arbitrary constants.

A descriptive differential invariant is a set of functions of x, p, the Π's, and the partial derivatives of the Π's up to any finite order:

$$F^{a_1 \cdots a_r}_{b_1 \cdots b_s}\left[x, p, \Pi, \frac{\partial \Pi}{\partial x}, \frac{\partial \Pi}{\partial p}, \cdots\right], \tag{7.6}$$

which are related to the same functions formed for $\bar{x}, \bar{p}, \bar{\Pi}$ and the derivatives of $\bar{\Pi}$:

$$F^{a_1 \cdots a_r}_{b_1 \cdots b_s}\left[\bar{x}, \bar{p}, \bar{\Pi}, \frac{\partial \bar{\Pi}}{\partial \bar{x}}, \frac{\partial \bar{\Pi}}{\partial \bar{p}}, \cdots\right], \tag{7.7}$$

by any prescribed law (7.1). If the law (7.1) is the tensor law, then we have a *descriptive tensor.*

The justification of the term *descriptive invariant* for (7.6) is as follows. Observe first that by § 4 the Γ's are determined when a parameterization is assigned to the paths, and conversely, when the Γ's are given, a parameterization of the paths is determined up to linear transformation. Now, a set of functions of the form (7.4) can be written in the form (7.6) when and only when these functions are unaffected by an arbitrary change of parameter on the paths. For if the functions (7.4) can be given the form (7.6), they are unaffected by change of parameter, since the Π's have this property. Conversely, if a set of functions (7.4) is independent of the parameterization of the paths, we may choose in particular the canonical parameter, for which the Γ's are equal to the Π's, and the set (7.4) then takes the form (7.6).

Finally, the components of a metric differential invariant have the form

$$F^{a_1 \cdots a_r}_{b_1 \cdots b_s}\left[x,\ p,\ g,\ \frac{\partial g}{\partial x},\ \frac{\partial g}{\partial p},\ \cdots,\ C,\ \frac{\partial C}{\partial x},\ \frac{\partial C}{\partial p},\ \cdots\right] \tag{7.8}$$

where g is the tensor of the metric, and C is the curvature vector.

Similarly, we may define affine, descriptive, metric, differential parameters by allowing the $F^{a_1 \cdots a_r}_{b_1 \cdots b_s}$ to involve arbitrary functions $\varphi(x, p)$ together with their partial derivatives up to any finite order.

An example of an affine tensor is H^i_{jkl}. Another is the Berwald curvature tensor:*

$$B^i_{jkl} = \left(\frac{\partial \Gamma^i_{jk}}{\partial x^l} + p^b\, \Gamma^a_{bl}\, H^i_{jka} - \Gamma^i_{al}\, \Gamma^a_{jk}\right) - \left(\frac{\partial \Gamma^i_{jl}}{\partial x^k} + p^b\, \Gamma^a_{bk}\, H^i_{jla} - \Gamma^i_{ak}\, \Gamma^a_{jl}\right). \tag{7.9}$$

The conditions

$$H^i_{jkl} = 0, \qquad B^i_{jkl} = 0 \tag{7.10}$$

are necessary and sufficient in order that it be possible by transformation of coördinates to reduce a given affine space of paths to the form $d^2x^i/dt^2 = 0$, characteristic of the straight lines, linearly parameterized, of a euclidean space. Or we may say that (7.10) is necessary and sufficient for the existence of a coördinate system in which the components Γ^i_{jk} of affine connection are all zero.

An example of a descriptive tensor, besides $\mathfrak{H}^i_{jkl}$, is the generalized Weyl tensor

$$\mathfrak{W}^i_{jkl} = \mathfrak{B}^i_{jkl} - \frac{\delta^i_k}{N-1}\mathfrak{B}^a_{jal} - \frac{\delta^j_l}{N-1}\mathfrak{B}^a_{jka}, \tag{7.11}$$

* (A), p. 215; (B), p. 46. Berwald uses the notation K^i_{jkl} instead of B^i_{jkl}.

where $\mathfrak{B}^i_{jkl}$ denotes B^i_{jkl} formed with the Π's in place of the Γ's. The proof of the tensor character of (7.11) is identical with that given for the restricted theory by J. M. Thomas,* being based on the law of transformation (5.9) of the Π's.

For $N>2$, the conditions

$$\mathfrak{H}^i_{jkl} = 0, \qquad \mathfrak{W}^i_{jkl} = 0 \tag{7.12}$$

are necessary and sufficient in order that a given system of paths be *linear*, i. e., equivalent by point transformation to the straight lines of a flat projective space. For $\mathfrak{H}^i_{jkl} = 0$ implies that the system of paths is of the restricted type, and then $\mathfrak{W}^i_{jkl} = 0$ secures linearity.‡ Conversely, if the system of paths is linear, then there exists a coördinate system together with a parameterization such that the differential equations of the paths have the form $d^2x^i/dt^2 = 0$. For this coördinate system and parameterization the equations (7.12) are obviously satisfied; hence since these equations have both descriptive and tensor character, they are satisfied in every coördinate system and parameterization.

8. **Affine normal coördinates.** Normal coördinates were introduced into the restricted affine geometry of paths by Veblen.† Suppose O is a fixed point, called the origin, in a restricted affine space of paths (0.3); then normal coördinates relative to the coördinate system (x) and origin O are defined as coördinates y, analytic functions of x, such that the paths through the origin have linear equations $y^i = a^i t + b^i$, where y vanishes at O. These conditions determine y up to homogeneous linear transformation; y is completely fixed by the further stipulation that at O, $dx = dy$, or $\partial x^i/\partial y^j = \partial y^i/\partial x^j = \delta^i_j$.

Let us try to introduce coördinates y with the same properties into a general affine space of paths

$$\frac{d^2x^i}{dt^2} = \underset{2}{H^i}(x, p). \tag{8.1}$$

The paths through the origin are required to have linear equations

$$y^i = p^i_0 (t - t_0), \tag{8.2}$$

where O corresponds to $t = t_0$, and $p^i_0 = (dy^i/dt)_{t_0} = (dx^i/dt)_{t_0}$.

---

* *Note on the projective geometry of paths*, Proc. Nat. Acad. Sci., vol. 11 (1925), pp. 207–209.

† *Normal coördinates for the geometry of paths*, Proc. Nat. Acad. Sci., vol. 8 (1922), pp. 192–197.

‡ H. Weyl, *Zur Infinitesimalgeometrie*, Gött. Nach., 1921, pp. 105, 106.

We need to find the Taylor series for the path determined by the initial conditions: for $t = t_0$, $x = x_0$ (coördinates of O), $dx/dt = p_0$. To this end we observe that

$$\frac{d^m x^i}{dt^m} = \underset{m}{H^i}(x, p), \tag{8.3}$$

the last symbol representing a function homogeneous of degree m in p. (8.3) results by mathematical induction from the fact that the operator d/dt applied to any function of x, p is equivalent to

$$p^j \frac{\partial}{\partial x^j} + \underset{2}{H^j} \frac{\partial}{\partial p^j},$$

and this has the effect of converting a function homogeneous of degree r in p into one homogeneous of degree $r+1$.

The Taylor expansion for the path determined by the stated initial conditions is therefore

$$x^i = x_0^i + p_0^i (t - t_0) + \sum_{m=2}^{\infty} \frac{1}{m!} \underset{m}{H^i}(x_0, p_0)\,(t - t_0)^m. \tag{8.4}$$

By the properties of homogeneous functions and (8.2), this becomes

$$x^i = x_0^i + \overset{\cdot}{y}^i + \sum_{m=2}^{\infty} \frac{1}{m!} \underset{m}{H^i}(x_0, y). \tag{8.5}$$

The transformation (8.5) defines the new coördinates y. If the $\underset{2}{H^i}$ are homogeneous quadratic polynomials, the $\underset{m}{H^i}$ are homogeneous polynomials of degree m, and the transformation from x to y defined by (8.5) is analytic at the origin. But in any other case (8.5) defines a non-analytic transformation. In fact, the second derivatives $\partial^2 x^i / \partial y^j \partial y^k$ are indeterminate at the origin, being $= \Gamma^i_{jk}(x_0, y)$, functions of the ratios of the y's, whose limiting values as we approach the origin depend effectively, in the general case, on the direction of approach.

Hence the following theorem, to which the special importance of systems of paths of the restricted type is largely due:

Let O denote an arbitrarily fixed point in an affine space of paths referred to a coördinate system (x); then when and only when the space of paths is of the restricted type does there exist, for every point O, a system of coördinates $(y)_O$ such that (1) *the transformation from x to y is analytic at O,* (2) *the paths through O have linear equations* (8.2).

Passing by this negative result, we proceed to define normal coördinates in a general affine space of paths relative to a fixed *point and direction*

x_0, p_0. These normal coördinates will no longer have the linear property (8.2), but other important properties of normal coördinates will be preserved.

We define the *osculating homogeneous polynomial in the direction* p_0 to the homogeneous function $\underset{m}{H^i}(x_0, y)$ to be

$$\underset{m}{P^i}(x_0, p_0; y) = \Gamma^i_{a_1 a_2 \cdots a_m} y^{a_1} y^{a_2} \cdots y^{a_m}, \tag{8.6}$$

where

$$\Gamma^i_{a_1 a_2 \cdots a_m} = \frac{1}{m!} \frac{\partial^m \underset{m}{H^i}(x, p)}{\partial p^{a_1} \partial p^{a_2} \cdots \partial p^{a_m}} \Bigg]_{x=x_0,\, p=p_0}. \tag{8.7}$$

The osculating analytic function in the direction p_0 to the (in general) non-analytic function (8.5) will mean the analytic function obtained by replacing each $\underset{m}{H^i}(x_0, y)$ by its corresponding $\underset{m}{P^i}(x_0, p_0; y)$:

$$x^i = x_0^i + y^i + \sum_{m=2}^{\infty} \frac{1}{m!} \Gamma^i_{a_1 a_2 \cdots a_m} y^{a_1} y^{a_2} \cdots y^{a_m}. \tag{8.8}$$

The multiple power series (8.8) is convergent. For on our making the substitution $y^i = p_0^i (t - t_0)$, (8.8) becomes (8.4) if we take account of Euler's theorem:

$$\frac{1}{m!} \frac{\partial^m \underset{m}{H^i}(x, p)}{\partial p^{a_1} \partial p^{a_2} \cdots \partial p^{a_m}} p^{a_1} p^{a_2} \cdots p^{a_m} = \underset{m}{H^i}(x, p). \tag{8.9}$$

Since (8.4) is convergent for sufficiently small values of $t - t_0$, (8.8) is convergent for sufficiently small values of y having the mutual ratios p_0. Therefore, by well-known theorems on power series, (8.8) is convergent for all sufficiently small values of y.

The coördinates y defined by the analytic transformation (8.8) are the normal coördinates relative to the coördinate system (x) and the point-direction x_0, p_0.

An alternative method of defining the same normal coördinates is the following, suggested to the writer by T. Y. Thomas. Write down the system of differential equations

$$\begin{aligned} \frac{d^2 x^i}{dt^2} &= \Gamma^i_{jk}(x, p) \frac{dx^j}{dt} \frac{dx^k}{dt}, \\ \frac{dp^i}{dt} &= \Gamma^i_{jk}(x, p) \frac{dx^j}{dt} p^k, \end{aligned} \tag{8.10}$$

to be considered as equations for x, p as functions of t. Suppose these equations solved under the initial conditions: for $t = t_0$, $x = x_0$, $p = p_0$,

$dx/dt = \xi$. The Taylor expansion for x can be calculated to have the form

$$x^i = x_0^i + \xi^i (t - t_0) + \frac{1}{2!} \Gamma^i_{jk}(x_0, p_0)\, \xi^j \xi^k (t - t_0)^2 + \frac{1}{3!} \Gamma^i_{jkl}(x_0, p_0)\, \xi^j \xi^k \xi^l (t - t_0)^3 + \cdots, \tag{8.11}$$

where Γ^i_{jk}, Γ^i_{jkl}, $\cdots$ have the values (8.7). Writing in (8.10)

$$y^i = \xi^i (t - t_0), \tag{8.12}$$

we then get the same equations for defining normal coördinates y which were obtained before as (8.8), viz.,

$$x^i = x_0^i + y^i + \frac{1}{2!} \Gamma^i_{jk}(x_0, p_0) y^j y^k + \frac{1}{3!} \Gamma^i_{jkl}(x_0, p_0) y^j y^k y^l + \cdots. \tag{8.13}$$

By means of this second definition it is easy to prove the important fact that when the coördinates x undergo arbitrary analytic transformation, the corresponding normal coördinates y undergo homogeneous linear transformation. For from $\bar{x}^i = \bar{x}^i(x)$, it follows that $d\bar{x}^i/dt = (\partial \bar{x}^i/\partial x^j)(dx^j/dt)$, or $\bar{\xi}^i = (\partial \bar{x}^i/\partial x^j)_0 \xi^j$; hence by (8.12)

$$\bar{y}^i = \left(\frac{\partial \bar{x}^i}{\partial x^j}\right)_0 y^j, \tag{8.14}$$

where the coefficients $(\partial \bar{x}^i/\partial x^j)_0$ are constants.

9. Differentiation of a tensor. The covariant derivative with respect to a general affine connection $\Gamma^i_{jk}(x, p)$ of a tensor whose components are functions of x, p has been defined by Berwald.* The idea is contained also in the generalization of Levi-Civita's parallel propagation of a vector given by Synge,† and in the "θ-process" of J. H. Taylor.‡

By applying the operator $\partial/\partial x^e$ to the tensor law

$$\frac{\partial \bar{x}^{\beta_1}}{\partial x^{b_1}} \cdots \frac{\partial \bar{x}^{\beta_s}}{\partial x^{b_s}} \bar{T}^{a_1 \cdots a_r}_{\beta_1 \cdots \beta_s}(\bar{x}, \bar{p}) = \frac{\partial \bar{x}^{a_1}}{\partial x^{\alpha_1}} \cdots \frac{\partial \bar{x}^{a_r}}{\partial x^{\alpha_r}} T^{\alpha_1 \cdots \alpha_r}_{b_1 \cdots b_s}(x, p), \tag{9.1}$$

and substituting for the second partial derivatives $\partial^2 \bar{x}^i/\partial x^j \partial x^k$ by means of (4.15), we find that

* (B), p. 51; (C), p. 198.
† Loc. cit., § 4.
‡ Loc. cit., § 3.

$$T^{a_1\cdots a_r}_{b_1\cdots b_s,c} = \frac{\partial T^{a_1\cdots a_r}_{b_1\cdots b_s}}{\partial x^c} + p^m \Gamma^l_{mc} \frac{\partial T^{a_1\cdots a_r}_{b_1\cdots b_s}}{\partial p^l} - \sum_{\varrho=1}^{r} \Gamma^{a_\varrho}_{\alpha c} T^{a_1\cdots a_{\varrho-1}\alpha a_{\varrho+1}\cdots a_r}_{b_1\cdots b_s} + \sum_{\sigma=1}^{s} \Gamma^{\beta}_{b_\sigma c} T^{a_1\cdots a_r}_{b_1\cdots b_{\sigma-1}\beta b_{\sigma+1}\cdots b_s} \tag{9.2}$$

are the components of a tensor. (9.2) is called the *covariant derivative of T with respect to x.* The operation of covariant differentiation may be repeated any number of times; thus $T^{a_1\cdots a_r}_{b_1\cdots b_s,c_1,c_2,c_3}$ is the notation for the third covariant derivative of T with respect to x.

Applying the operator $\partial/\partial p^d$ to (9.1), we get an equation which expresses that $\partial T^{a_1\cdots a_r}_{b_1\cdots b_s}/\partial p^d$, which we shall write

$$T^{a_1\cdots a_r}_{b_1\cdots b_s/d}, \tag{9.3}$$

are the components of a tensor.* (9.3) is called the *derivative of T with respect to p.* Differentiation of a tensor with respect to p may be carried out successively any number of times or combined in any order with covariant differentiation as to x. Thus we have the tensor

$$T^{a_1\cdots a_r}_{b_1\cdots b_s,c_1,c_2/d_1d_2d_3,e_3/e_1e_2}.$$

This is evidently symmetric in the indices d and the indices e of each separate group of differentiations with respect to p.

10. **Extension.** More convenient for many purposes than the process of differentiation of a tensor is that of *extension* of a tensor.†

Let $T^{a_1\cdots a_r}_{b_1\cdots b_s}(x, p)$ be the components of a tensor T in the coördinate system (x). Suppose the representation of T in the system of normal coördinates y determined by (x) together with an arbitrarily fixed point-direction x, p to be $t^{a_1\cdots a_r}_{b_1\cdots b_s}(y, q)$, where $q = dy/dt$; i. e.,

$$t^{a_1\cdots a_r}_{b_1\cdots b_s} = T^{\alpha_1\cdots\alpha_r}_{\beta_1\cdots\beta_s} \frac{\partial y^{a_1}}{\partial x^{\alpha_1}} \cdots \frac{\partial y^{a_r}}{\partial x^{\alpha_r}} \frac{\partial x^{\beta_1}}{\partial y^{b_1}} \cdots \frac{\partial x^{\beta_s}}{\partial y^{b_s}}. \tag{10.1}$$

Then the quantities

$$\left.\frac{\partial^{m+n} t^{a_1\cdots a_r}_{b_1\cdots b_s}(y, q)}{\partial q^{c_1}\cdots\partial q^{c_m}\partial y^{d_1}\cdots\partial y^{d_n}}\right]_{y=0,\ q=p}, \tag{10.2}$$

* Berwald, (B), p. 52; (C), p. 200.

† For the restricted theory the idea of extension is due to Veblen and T. Y. Thomas, loc. cit., § 11.

which are functions of x, p, are the components of a tensor, to be called the *m, nth extension of T as to p, x*. The proof of the tensor character of (10.2) is identical with the proof of the corresponding fact in the restricted theory,* being based on the linear transformation property of normal coördinates established at the end of § 8.

If (10.2) contains no differentiations as to q, we have the nth extension of T as to x, for which we shall use the notation

$$\frac{\partial^n t^{a_1 \cdots a_r}_{b_1 \cdots b_s}(y, q)}{\partial y^{d_1} \cdots \partial y^{d_n}}\Bigg]_{y=0,\ q=p} = T^{a_1 \cdots a_r}_{b_1 \cdots b_s d_1 \cdots d_n}(x, p) \tag{10.3}$$

The first extension of a tensor as to x is the same as its covariant derivative. For by (8.8) we have

$$\frac{\partial x^i}{\partial y^j}\Bigg]_0 = \delta^i_j, \qquad \frac{\partial^2 x^i}{\partial y^j \partial y^k}\Bigg]_0 = \Gamma^i_{jk}(x_0, p_0); \tag{10.4}$$

hence by (4.15), if $C^i_{jk}(y, q)$ denote the representation of Γ^i_{jk} in normal coördinates relative to x_0, p_0, we have $C^i_{jk}(0, p_0) = 0$—from this, by (9.2), the statement just made follows.

It may be shown that the process of forming the m, nth extension of a tensor as to p, x is equivalent to taking the mth derivative of the tensor as to p followed by taking the nth extension as to x; i. e., combining the notation (10.3) and (9.3), (10.2) is equal to

$$T^{a_1 \cdots a_r}_{b_1 \cdots b_s / c_1 \cdots c_m, d_1 \cdots d_n}. \tag{10.5}$$

For, by applying the operator $\partial^m / \partial q^{c_1} \cdots \partial q^{c_m}$ to (10.1), we get

$$\frac{\partial^m t^{a_1 \cdots a_r}_{b_1 \cdots b_s}}{\partial q^{c_1} \cdots \partial q^{c_m}} = T^{a_1 \cdots a_r}_{\beta_1 \cdots \beta_s / \gamma_1 \cdots \gamma_m} \frac{\partial y^{a_1}}{\partial x^{\alpha_1}} \cdots \frac{\partial y^{a_r}}{\partial x^{\alpha_r}} \frac{\partial x^{\beta_1}}{\partial y^{b_1}} \cdots \frac{\partial x^{\beta_s}}{\partial y^{b_s}} \frac{\partial x^{\gamma_1}}{\partial y^{c_1}} \cdots \frac{\partial x^{\gamma_m}}{\partial y^{c_m}}, \tag{10.6}$$

which expresses that $\partial^m t^{a_1 \cdots a_r}_{b_1 \cdots b_s} / \partial q^{c_1} \cdots \partial q^{c_m}$ is the representation in normal coördinates of $T^{a_1 \cdots a_r}_{b_1 \cdots b_s / c_1 \cdots c_m}$, or, as we write it,

$$\frac{\partial^m t^{a_1 \cdots a_r}_{b_1 \cdots b_s}}{\partial q^{c_1} \cdots \partial q^{c_m}} = t^{a_1 \cdots a_r}_{b_1 \cdots b_s / c_1 \cdots c_m} \tag{10.7}$$

—in other words, the operations of (1) representation in normal coördinates, (2) differentiation with respect to p (or q), are commutative. The state-

* Loc. cit., §§ 10, 11.

ment at the beginning of this paragraph now follows by operating on (10.7) with $\partial^n/\partial y^{d_1}\cdots\partial y^{d_n}$ and evaluating for $y = 0$, $q = p$.

11. **Normal tensors.** The process of extension may be applied to any invariant as well as to a tensor. Denoting by $C^i_{jk}(y, q)$ the representation of $\Gamma^i_{jk}(x, p)$ in normal coördinates relative to (x), x, p, we can show exactly as in the restricted theory* that the extensions of the affine connection,

$$(11.1)\qquad \left.\frac{\partial^{m+n} C^i_{jk}(y, q)}{\partial q^{c_1}\cdots\partial q^{c_m}\,\partial y^{d_1}\cdots\partial y^{d_n}}\right]_{y=0,\,q=p} = H^i_{jk\,c_1\cdots c_m,\,d_1\cdots d_n},$$

are tensors. The proof depends on the fact that in the linear transformation of normal coördinates from y and $\bar{y}$ the affine connection $C^i_{jk}(y, q)$ behaves like a tensor, since in (4.15) (written in y, $\bar{y}$) the second derivatives $\partial^2\bar{y}^i/\partial y^j\,\partial y^k$ vanish. The tensors (11.1) will be called *normal tensors*.† They are symmetric in the indices $j, k, c_1, \cdots, c_m$, on account of $\partial C^i_{jk}/\partial q^c = \partial C^i_{jc}/\partial q^k$, and likewise in the indices $d_1, \cdots, d_n$.

We may also define generalized normal tensors,‡ derived from the higher Γ's of § 8. Let $C^i_{jk\cdots l}(y, q)$ be the representation in normal coördinates relative to the point-direction x, p of $\Gamma^i_{jk\cdots l}$. Then

$$(11.2)\qquad \left.\frac{\partial^{m+n} C^i_{jk\cdots l}}{\partial q^{c_1}\cdots\partial q^{c_m}\,\partial y^{d_1}\cdots\partial y^{d_n}}\right]_{y=0,\,q=p}$$

are the components of a tensor, which we shall denote by $H^i_{(jk\cdots l)\,c_1\cdots c_m,\,d_1\cdots d_n}$.

The expressions for the successive covariant derivatives of a tensor T in terms of the extensions of T and the normal tensors are more complicated than in the restricted theory;§ for instance:

$$(11.3)\qquad T_{ij,\,k,\,l} = T_{ij,\,kl} + p^\beta H^\alpha_{\beta k,\,l}\,T_{ij\alpha} + H^\alpha_{ik,\,l}\,T_{\alpha j} + H^\alpha_{jk,\,l}\,T_{i\alpha}.$$

This is derived by differentiating the formula for $T_{ij,\,k}$, expressed in normal coördinates, with respect to y^l and evaluating for the original point-direction.

---

* Veblen and T. Y. Thomas, § 9.

† In the restricted theory all normal tensors which involve differentiation as to q vanish. The non-vanishing ones, $H^i_{jk,\,d_1\ldots d_n}$ are denoted by $A^i_{jk\,d_1\ldots d_n}$.

‡ Cf. loc. cit., § 13.

§ Loc. cit., § 12.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

ON A GEOMETRICAL THEORY OF CONTINUOUS GROUPS.

II. EUCLIDEAN AND HYPERBOLIC GROUPS OF THREE-DIMENSIONAL SPACE.*

By B. de Kerékjártó, Szeged (Hungary).

1. The purpose of the present paper is to give a simple topological characterization of the groups of three-dimensional euclidean and hyperbolic geometries. For the case of two dimensions the problem has been solved by Hilbert,† and we make use here of his results and of some of his methods. Besides the axioms of Hilbert we assume also the topological content of an axiom used by Lie.‡ A subsequent paper will deal with the problem of reducing the present system of axioms to a less restrictive one which will be the direct analogue to the system of Hilbert.

2. We consider a *group* Γ of topological transformations of three-dimensional cartesian space into itself with invariant indicatrix, about which we make the following assumptions:

(*a*) The transformations of the group form a *closed system* (in the sense of Hilbert); this means that if A, B, C, D and A', B', C', D' are arbitrary points in the space such that there are transformations of the group carrying points arbitrarily near to A, B, C, D into points arbitrarily near to A', B', C', D' respectively, there is a transformation in the group carrying A, B, C, D into A', B', C', D'.

(*b*) If A is an arbitrary point of space, then under the subgroup of the given group which consists of the transformations leaving A invariant every other point describes a *closed surface*. We call such a surface a *pseudosphere* around A.

(*c*) If B is another arbitrary point, every point into which B is carried by the transformations leaving A invariant—except a finite number of them—is carried into infinitely many points by the subgroup which leaves both A and B invariant.

3. We take an arbitrary point O of space and consider the pseudospheres around O. They form a family of closed surfaces such that through every point of space except O there passes one and only one surface of the

* Received July 9, 1927. For the first communication, see Annals of Math. (2), vol. 27 (1925), pp. 105–117. The second communication has been presented to the Hungarian Academy of Sciences at the meeting of Nov. 7, 1927.

† Hilbert, Über die Grundlagen der Geometrie, Math. Annalen, vol. 56 (1902), pp. 381–422; reprinted in Hilbert, Grundlagen der Geometrie, 4. ed. Leipzig and Berlin, 1913, Anhang IV, pp. 163–218.

‡ Lie-Engel, Theorie der Transformationsgruppen, vol. III (1893), p. 506 et seq.

family. Hence it follows that every surface of the family is homoeomorphic with an ordinary sphere, that each of them contains O in its interior, and that of any two of them one lies inside of the other one.* We set the pseudospheres σ_O around O into correspondence with the positive real values r in such a way that if $r<r'$ then σ_O^r lies inside of $\sigma_O^{r'}$; this correspondence is then continuous in that sense that the points of σ_O^r and $\sigma_O^{r'}$ all lie arbitrarily near to $\sigma_O^{r'}$ and σ_O^r respectively if $|r-r'|$ is sufficiently small.

4. We consider a pseudosphere σ_O around O; we denote by Γ_O the subgroup consisting of those transformations which leave O invariant. Every transformation of Γ_O furnishes a topological transformation of σ_O into itself; these transformations of σ_O form a group which essentially satisfies the system of axioms of Hilbert.† In particular it can be shown by the methods of Hilbert‡ *that the subgroup Γ_{OP} of Γ_O which leaves the point P of σ_O invariant is equivalent with regard to σ_O with the group of rotations of a sphere around one of its axes;* we call this group a *cyclic group of rotations of σ_O around OP.* Under this group a certain uniquely determined point P' of σ_O remains invariant besides P; the points P and P' will be denoted as a couple of *diametrically opposite* or *antipodal* points of σ_O. Diametrically opposite points of σ_O are carried by every transformation of Γ_O into diametrically opposite points; indeed if t is an arbitrary transformation of Γ_O which carries P into Q, we have:

$$\Gamma_{OQ} = t^{-1}\Gamma_{OP}t.$$

5. If t is an arbitrary transformation of Γ_O it has at least one invariant point P on σ_O, by the theorem of Brouwer;§ the cyclic group Γ_{OP} contains then the transformation t. It follows that every transformation of Γ_O belongs to some cyclic subgroup of Γ_Q.

Under the cyclic subgroup Γ_{OP} the two opposite points P and P' are invariant and every other point of σ_O describes on σ_O a simple closed curve; these curves will be called *pseudocircles* on σ_O around PP'.

Under a transformation t of Γ_O which carries P into a point Q every pseudocircle around PP' is carried into a pseudocircle around QQ'; this is a direct consequence of the above relation: $\Gamma_{OQ} = t^{-1}\Gamma_{OP}t$.

---

* This proposition will be proved in a paper of the series „Sur les familles de surfaces et de courbes" soon to be published in the Acta Literarum ac Scientiarum of the University of Szeged.

† Hilbert, Grundlagen der Geometrie, pp. 167–169; the second axiom holds in our case for a neighborhood of the invariant point.

‡ Hilbert, l. c., pp. 173 et seq.

§ Brouwer, Math. Annalen, vol. 71, pp. 114, 326.

6. Under a transformation t of Γ_0 which carries P into its opposite point P' the family of the pseudocircles on σ_0 around PP' is carried into itself. Hence it follows that there is one and only one among these circles which is carried into itself; we denote this one by $\varkappa$ and call it the *great circle* around PP'. On $\varkappa$ there are precisely two points Q and Q' left invariant by this transformation, which are necessarily two diametrically opposite points. The transformation t is then a rotation around QQ'; its square carries each of the points P, P', Q, Q' into itself so that t^2 is the identity. We call t the *half-rotation* around QQ' and denote it by $t^{1/2}_{QQ'}$. The two arcs of $\varkappa$ determined by Q and Q' are transformed into one another by $t^{1/2}_{QQ'}$; there is precisely one couple of diametrically opposite points R, R' on $\varkappa$ such that R and R' are interchanged by $t^{1/2}_{QQ'}$. The half-rotation around any couple of diametrically opposite points of $\varkappa$ carries P into P'. By the half-rotation $t^{1/2}_{PP'}$ around PP' every point on $\varkappa$ is carried into its diametrically opposite point. We call the couples of diametrically opposite points P, P' and Q, Q' *orthogonal* to one another if the half-rotation around one of them interchanges the two points of the other couple, and thus vice versa. All the couples of opposite points which are orthogonal to PP' lie on the great circle around PP'.

From these propositions it follows, further, that a pseudocircle such that the point diametrically opposite to each point of the pseudocircle also lies on the circle is necessarily a great circle on σ_0.

7. Let $\varkappa$ be a great circle, PP' its centers, QQ' a couple of opposite points on $\varkappa$. By the rotations around QQ' the great circle $\varkappa$ is carried into great circles whose family covers the whole sphere σ_0. Hence it follows that for any two points of σ_0 there is at least one great circle passing through both of them.

Let $\varkappa$ and $\varkappa'$ be two arbitrary great circles on σ_0 with the centers PP' and QQ' respectively; let $\bar{\varkappa}$ be a great circle passing through P and Q; denote the centers of $\bar{\varkappa}$ by RR'. From § 6 it follows that $\varkappa$ and $\varkappa'$ pass through R and R' so that they have at least two points common. They cannot have any other common points unless they are identical. Indeed if PP' and QQ' are different, and if SS' is another couple of opposite points belonging to $\varkappa$ and $\varkappa'$ the half-rotation $t^{1/2}_{RR'}$ interchanges P and P', Q and Q'. The same holds for the half-rotation $t^{1/2}_{SS'}$ around SS', so that the transformation $t^{1/2}_{RR'}\, t^{1/2}_{SS'}$ would have four invariant points P, P', Q, Q'. This transformation would be the identity which would give a contradiction. It follows that:

Two arbitrary great circles on σ_0 have precisely two common points which are necessarily diametrically opposite to one another.

We observe that a great circle subdivides the pseudosphere into two

regions such that the point diametrically opposite to any point of one of these regions belongs to the other region. Consider indeed a great circle $\varkappa$ around PP' and a great circle $\varkappa'$ through PP'. $\varkappa$ and $\varkappa'$ have two common points Q and Q'. The points on that arc PQ of $\varkappa'$ which does not contain P' and Q' have their opposites on that arc $P'Q'$ of $\varkappa'$ which does not contain P and Q, and so on. From this the above proposition follows immediately.

8. We have obtained the result that to any two points of σ_O which are not diametrically opposite there is one and only one great circle on σ_O passing through both of them. From this we deduce the following proposition.

If P and Q are two arbitrary points on σ_O sufficiently near to each other there is a transformation of Γ_O which carries P into Q and which changes all the points of σ_O by as little as we please.

Consider indeed the great circle on σ_O through P and Q, and denote its centers by RR'; a rotation around RR' which carries P into Q differs very little on σ_O from the identity if P and Q are sufficiently near to each other; this follows immediately from axiom (a), to the effect that the transformations of the group form a closed system.

9. From § 4 it follows that a transformation of the group which leaves O and two non opposite points of the pseudosphere σ_O invariant is the identity on σ_O. By considering the pseudospheres around a point P of σ_O we find that the same transformation is the identity for every other pseudosphere σ'_O around O; that is to say the identity for the whole space.

If we take a transformation of the pseudosphere σ_O into itself such that its n-th power is the identity on σ_O, the transformation is periodic of period n for the whole space. Hence it follows that the subgroup Γ_{OP} which leaves O and a point P of σ_O invariant is a cyclic group for the whole space e. g. a cyclic group for every pseudosphere σ'_O around O.

To every couple of diametrically opposite points PP' on σ_O corresponds in this way a uniquely determined couple of opposite points on every other pseudosphere σ'_O around O namely the invariant points of σ'_O for the cyclic group Γ_{OP}. The set of points which are invariant for Γ_{OP} will be called a *pseudoline* passing through O and P and denoted by λ_{OP}.

As the points O and P are two arbitrary points of the space we can state:

Any two different points of the space determine one and only one pseudoline passing through both of them.

From the above remarks it follows also that:

Any two different points of a pseudoline determine this pseudoline.

10. *A pseudoline is a simple continuous open line,* that is, a closed point-set which is a continuous one-to-one image of an infinite straight

line. By the results of § 3 there is a continuous one-to-one correspondence between the set of the pseudospheres σ_O around O and the set of positive real numbers r. On every pseudosphere σ_O there are two points of the pseudoline λ_{OP}; these points together form the set of invariant points of a topological transformation of the space into itself so that their set is closed. Consequently if the pseudospheres $\sigma_O^{(1)}$, $\sigma_O^{(2)}$, $\cdots$ converge to the pseudosphere $\sigma_O^{(\omega)}$ the points of λ_{OP} lying respectively on $\sigma_O^{(1)}$, $\sigma_O^{(2)}$, $\cdots$ converge to the points of the same pseudoline lying on $\sigma_O^{(\omega)}$. From this follows the existence of a topological correspondence between the real numbers and the points of the pseudoline.

11. *A pseudoline is transformed by any transformation of the given group into a pseudoline.*

The pseudoline λ_{OP} through O and P is the set of invariant points of the group Γ_{OP}. The image of λ_{OP} by a transformation t which carries O and P into the points Q and R respectively is the set of invariant points of the group Γ_{QR} (for $\Gamma_{QR} = t^{-1}\Gamma_{OP}t$); that is, the pseudoline λ_{QR}.

We understand by *segment* PQ of a pseudoline λ_{PQ} those points of λ_{PQ} which lie between P and Q (in the sense of a continuous one-to-one correspondence with an infinite straight line).

Two segments PQ and RS on the same or on two different pseudolines are called *congruent* if there is a transformation in the given group Γ which carries the segment PQ into the segment RS.

A segment PQ cannot be congruent to a part of itself. It is enough to show that PQ cannot be congruent to a segment PR if R lies on the segment PQ and is different from Q. Indeed if the segments PQ and PR are congruent there is a transformation in the group which carries P into itself and the point Q into the point R; in other words, there is a rotation around P which carries Q into R. By this rotation the pseudosphere around P passing through Q is invariant, so that Q and R both lie on this pseudosphere. But this is a contradiction, for the segment PQ has only the point Q on the pseudosphere around P through Q.

Hence it follows that a segment PQ has a uniquely determined mid-point R such that the segments PR and RQ are congruent. (There is a sphere around R which passes through both of the points P and Q). It is obvious from the preceding results that by a transformation which carries the segment PQ into the segment $P'Q'$ the mid-point of the first segment is transformed into the mid-point of the second one.

12. *The pseudoline passing through O and P changes continuously by continuous variation of the point P.*

It is sufficient to show this for the case where P is varying on a pseudosphere σ_O. Let P be a fixed and P_1 a variable point on σ_O; we

want to show that all the points of the pseudoline λ_{OP_1} converge to the pseudoline λ_{OP} if P_1 converges to P. By the result of § 8 there is a neighborhood of the identity in the group of transformations of σ_O into itself (generated by Γ_O) such that its transformations carry the point P into all the points of a neighborhood of P on σ_O. From the condition (a) of § 2 combined with the results of § 9 it follows that the same transformations differ very little from the identity for any bounded region of the space. Hence a transformation of σ_O into itself which carries P into a neighboring point P_1 along the great circle through P and P_1 transforms the pseudoline λ_{OP} into a pseudoline λ_{OP_1} which lies in any finite region of the space arbitrarily near to λ_{OP} provided that P and P_1 lie sufficiently near to one another.

By a double application of this result we obtain:

The pseudoline λ_{OP} passing through O and P changes continuously by continuous variation of the points O and P.

13. Let λ_{OP} be the pseudoline determined by O and P and let λ_{OQ} be another pseudoline passing through the points O and Q. Suppose that P and Q lie on the same pseudosphere σ_O around O and denote on this their opposite points by P' and Q'. If the couples of points PP' and QQ' are orthogonal to each other (in the sense of § 6) we say that the *pseudolines λ_{OP} and λ_{OQ}* are *orthogonal* to each other. It is obvious that this definition is independent of the special choose of the pseudosphere σ_O. To recollect the meaning of this notion: λ_{OP} and λ_{OQ} are orthogonal to one another if and only if the half-rotation around $\lambda_{OP}(\lambda_{OQ})$ carries λ_{OQ} (λ_{OP}) into itself in such a way that the two half-lines of λ_{OQ} (λ_{OP}) determined by O are interchanged.

If λ_{OR} is any pseudoline into which a pseudoline λ_{OQ} orthogonal to λ_{OP} is carried by an arbitrary rotation around λ_{OP}, then λ_{OP} and λ_{OR} are orthogonal. All the lines through O which are orthogonal to λ_{OP} are obtained from one of them by rotations around λ_{OP}.

14. The set of points belonging to the pseudolines λ_{OR} which are orthogonal to λ_{OP} and pass through O forms by definition a *pseudoplane Π_O*. We call *O origin of the pseudoplane Π_O* and λ_{OP} the orthogonal axis of Π_O.

If O, R, S are three arbitrary points of space which do not belong to one and the same pseudoline, there is one and only one pseudoplane with origin O which passes through R and S. Consider indeed the pseudolines λ_{OR} and λ_{OS}; take a pseudosphere σ_O around O and denote by R_1, R_1', S_1, S_1' the points of intersection of λ_{OR} and of λ_{OS} with σ_O. Take the great circle on σ_O through R_1 and S_1 (this passes also through R_1' and S_1') denote its centers by P_1, P_1'. The pseudoplane Π_O with origin O and with the orthogonal axis λ_{OP_1} contains the points R and S and is the only

pseudoplane with origin O containing these points. We denote this pseudoplane determined by the origin O and by the two other points R and S by $\Pi_{O,RS}$.

15. From the definition of orthogonality it is clear that *two orthogonal lines are carried into orthogonal lines by any transformation of* Γ. Indeed let λ_{OP} and λ_{OQ} be two orthogonal lines, t a transformation of Γ which carries O into a point O_1 the points P, Q into the points P_1, Q_1. The rotations t_{OP} which leave O and P invariant are transformed by t into rotations $t^{-1} t_{OP} t = t_{O_1P_1}$ which leave O_1 and P_1 invariant. By the rotation $t_{O_1P_1}^{1/2} = t^{-1} t_{OP}^{1/2} t$ the point Q_1 is carried into Q_1' which is the image of Q' by t, where Q' denotes the image of Q by $t_{OP}^{1/2}$. Since a pseudoline is transformed by any transformation of Γ into a pseudoline (§ 11), it follows that $\lambda_{O_1P_1}$ and $\lambda_{O_1Q_1}$ are orthogonal.

From this it follows that *the image under a transformation* t *of* Γ *of a pseudoplane* Π_O *with origin* O *and with orthogonal axis* λ_{OP} *is a pseudoplane* Π_{O_1} *with origin* O_1 *and with orthogonal axis* $\lambda_{O_1P_1}$, where O_1 and P_1 denote the images of O and P, respectively, by the transformation t.

By a rotation around the orthogonal axis λ_{OP} the pseudoplane is, of course, carried into itself.

16. With the aid of the above results on pseudolines and pseudoplanes we want to deduce some propositions about pseudospheres.

If A *and* B *are two arbitrary non opposite points on a pseudosphere* σ_O *and* C *is the mid-point of the segment* AB, *then the pseudolines* λ_{AB} *and* λ_{OC} *are orthogonal.*

Take the great circle $\varkappa$ on σ_O through A and B and denote by C^* the mid-point of the arc AB of $\varkappa$ (this is by definition a point on $\varkappa$ such that there is a rotation around the axis of $\varkappa$ which carries A into C^* and C^* into B). By the half-rotation around OC^* the points A and B are carried into one another; the pseudoline λ_{AB} has for image $\lambda_{BA} \equiv \lambda_{AB}$, the segment AB is carried into the segment BA, so that its mid-point C is invariant. It follows that C lies on λ_{OC^*} and that the half-rotation around λ_{OC} carries λ_{AB} into itself, that is to say: λ_{AB} and λ_{OC} are orthogonal.

17. *A pseudosphere* σ_O *and a pseudoline* λ_{AB} *cannot have more than two points in common.*

Suppose A, B and C are three points of λ_{AB} on σ_O. From § 9 it follows that no two of them can be diametrically opposite points on σ_O. Denote by T and U the mid-points of the segments AB and BC respectively; obviously $T \neq U$. The pseudolines λ_{OT} and λ_{OU} are orthogonal to λ_{AB} by the result of the preceding paragraph; by the half-rotation around λ_{AB} they are carried into themselves. The image of O by this half-rotation is a point O' different from O for O does not belong to λ_{AB}. The two pseudo-

lines λ_{OT} and λ_{OU} would then have two common points O and O' and this is a contradiction (see § 9).

If A and B are two arbitrary points on a pseudosphere σ_O the segment AB of the pseudoline λ_{AB} lies completely inside of σ_O except its extremities.

Suppose the contrary; let C, D, E be three points of λ_{AB} lying outside of σ_O such that C lies between A and B; D lies on the left hand side of A; and E on the right of B. Determine around each of these three points a neighborhood which does not contain points of σ_O. On the other hand take two points A' and B' in the interior of σ_O sufficiently near to A and B so that the pseudoline $\lambda_{A'B'}$ has at least one point C', D', E' in each of the neighborhoods determined around C, D, E; this is possible in consequence of § 12. The points D', A', C', B', E' lie on $\lambda_{A'B'}$ in this linear order, any two consecutive ones are in two different regions determined by σ_O; hence it follows that between any two of them on $\lambda_{A'B'}$ there is at least one point of σ_O. Thus, $\lambda_{A'B'}$ would have at least four points on σ_O and this would be in contradiction to the above result.

18. *Let A and B be two arbitrary points of the space, C a point on the segment AB of the pseudoline λ_{AB}; denote by σ_A and σ_B the pseudospheres through C with center A and B respectively. σ_A and σ_B have no other point in common than C.*

It is clear that σ_A and σ_B have on λ_{AB} no other common point than C. Let C_1 be a common point of σ_A and σ_B which does not belong to λ_{AB}. On the pseudosphere σ_A take the pseudocircle around λ_{AB} passing through C_1; denote on this by C_2 the image of C_1 by the half-rotation $t_{AB}^{1/2}$; the mid-point D of the segment C_1C_2 is a point of λ_{AB} and the pseudoline λ_{C_1D} is orthogonal to λ_{AB} (§ 16). As the points C_1 and C_2 belong to σ_A by the result of § 17 all the points of the segment C_1C_2 belong to the interior of σ_A. The point D of this segment belongs then to that segment of λ_{AB} which lies inside of σ_A. By the same reasoning the point D belongs to that segment of λ_{AB} which lies inside of σ_B. But these two segments of λ_{AB} have no point in common for they are separated from each other by the point C. This is a contradiction.

19. We define an *antipodal transformation* α *with center* O as follows: Let P be any point of space, σ_O the pseudosphere around O passing through P; denote by P' the diametrically opposite point of P on σ_O. The image of P by the transformation α should be the point P'.

It is clear that α is a continuous one-to-one transformation of space into itself with the single invariant point O. The square of α is the identity.

Each pseudoline through O is carried by α into itself. We want to show that every pseudoline is carried by α into a pseudoline. First we remark that any couple of points (A, B) is carried by α into a congruent

couple (A', B'), this means that there is also a transformation of the given group Γ which carries A into A' and B into B'. Indeed, denote by σ_O and σ'_O the pseudospheres around O passing through A and B respectively; denote by λ_{OP} the pseudoline which is orthogonal to λ_{OA} and λ_{OB} and by $\varkappa$ and $\varkappa'$ respectively the great circles on σ_O and on σ'_O around λ_{OP}. By the half-rotation $t^{1/2}_{OP}$ around λ_{OP} the points A and B are carried into their diametrically opposite ones on σ_O and σ'_O respectively, that is to say, into the same points as by α. — Let C be an arbitrary point of the segment AB, C' its image by α, C^* its image by $t^{1/2}_{OP}$. C^* lies on the segment $A'B'$ and we have the congruences: $(A, C) \equiv (A', C')$, $(A, C) = (A', C^*)$, thus $(A', C') = (A', C^*)$; this means that the pseudosphere around A' passing through C^* passes also through C'. By the same reasoning the pseudosphere around B' passing through C^* passes through C'. But these two pseudospheres have only one common point in consequence of the result of § 18, thus $C' \equiv C^*$. Herewith is proved that the image of the segment AB by the transformation α is the segment $A'B'$.

20. From the result of the preceding paragraph we obtain immediately the following one:

If A and B are two non opposite points on a pseudosphere σ_O, all the pseudolines passing through O and a point C of the segment AB intersect the great circle on σ_O passing through A and B.

Take the antipodal transformation α with center O and denote by A', B' the images of A, B by α. Then take the half-rotation $t^{1/2}_{OP}$ around that axis λ_{OP} which is orthogonal to λ_{OA} and λ_{OB}. By the transformation $\alpha \cdot t^{1/2}_{OP}$ the points A, B are invariant; moreover, each point of the segment AB is invariant. Each point of the great circle $\varkappa$ passing through A and B is carried into itself by this transformation, and they are the only invariant points on σ_O. The pseudoline λ_{OC} is transformed into itself by this transformation, for O and C are invariant. Consequently its point of intersection C_1 with σ_O is invariant so that C_1 belongs to $\varkappa$.

From the last result it follows at once:

If A and B are two arbitrary points of the pseudolines λ_{OA} and λ_{OB}, and C denotes an arbitrary point of the pseudoline λ_{AB} then λ_{OA}, λ_{OB} and λ_{OC} are orthogonal to the same pseudoline λ_{OP}.

This is equivalent with the following proposition (cf. § 14):

If two points A and B lie in a pseudoplane all the points of λ_{AB} lie in the same pseudoplane:

21. Next we want to show:

The pseudoplanes $\Pi_{A,BC}$ and $\Pi_{B,AC}$ are identical.

Let λ_{AP} be the common orthogonal line to λ_{AB} and λ_{AC} and let λ_{BQ} be the orthogonal to λ_{BA} and λ_{BC}. If D is an arbitrary point on the pseudo-

line λ_{AC} then λ_{BD} is orthogonal to λ_{BQ} so that D lies in $\Pi_{B,AC}$. If E is an arbitrary point of $\Pi_{A,BC}$ consider on a pseudosphere σ_A around A the great circle $\varkappa$ with axis λ_{AP} and denote the points of intersection of λ_{AB}, λ_{AC}, λ_{AE} with $\varkappa$ by $B_1, B_1', C_1, C_1', E_1, E_1'$. Choose the notations in such a way that C_1 and E_1 lie on the same half-circle of $\varkappa$ determined by the points B_1 and B_1'.

If C_1 lies between B_1 and E_1 on this half-circle, take a variable point S of the segment BE and denote by S_1 that intersection of λ_{AS} with σ_A which lies on the preferred halfcircle of $\varkappa$. While the point S describes the segment BE from B to E the point S_1 describes continuously the arc $B_1C_1E_1$ of $\varkappa$. Hence it follows that there is a position where S_1 coincides with C_1, and then S lies on λ_{AC}. Our result is that the segment BE has at least one point S on λ_{AC}. The points B and S belong to the pseudoplane $\Pi_{B,AC}$, thus all the points of λ_{BS} and consequently also E belongs to $\Pi_{B,AC}$.

If on the arc of $\varkappa$ the point E_1 lies between B_1 and C_1, we show in the same way that the segment BC intersects the pseudoline λ_{AE}; the point of intersection S belongs to $\Pi_{B,AC}$, so that all the points of $\lambda_{AS} \equiv \lambda_{AE}$, e. g. the point E belongs to $\Pi_{B,AC}$. By this it is shown that all the points of $\Pi_{A,BC}$ belong to $\Pi_{B,AC}$. Changing the rôle of A and B we obtain the reciprocal of this proposition, and from both the identity of the two pseudoplanes.

22. If A', B', C' are three arbitrary points of the pseudoplane Π_{ABC} which do not belong to the same pseudoline then $\Pi_{A'B'C'} = \Pi_{ABC}$. Indeed, instead of B and C, we can take any two other points B' and C' of the same plane which do not belong to the same pseudoline through A, that is: $\Pi_{ABC} = \Pi_{AB'C'}$ (If A, B', C' lie on the same pseudoline, interchange the notations of A' and B'). Further: $\Pi_{AB'C'} = \Pi_{B'AC'}$ by the result of the preceding paragraph. Instead of A we can take any other point A' of the same plane which does not lie on the pseudoline through B' and C', thus: $\Pi_{B'AC'} = \Pi_{B'A'C'}$ and then $\Pi_{B'A'C'} = \Pi_{A'B'C'}$ so that finally we obtain: $\Pi_{ABC} = \Pi_{A'B'C'}$.

The obtained result can be stated as follows:

Any three points of a pseudoplane which do not belong to the same pseudoline determine this pseudoplane.

23. *If two pseudoplanes have a common point they have at least one more point in common.*

Let O be a common point of the two pseudoplanes; we can consider it by the result of § 21 as origin of each of these pseudoplanes. Take a pseudosphere σ_O around O; its intersection with one of the given pseudoplanes is a great circle $\varkappa$ on σ_O; its intersection with the other pseudoplane is another great circle $\varkappa'$ on σ_O. The great circles $\varkappa$ and $\varkappa'$ have

two common points on σ_0 (§ 7); they are common points of the given pseudoplanes.

24. We have obtained a system of points, lines, planes (which we have denoted as pseudolines and pseudoplanes) such that for them all the axioms of Hilbert* hold, except, of course, the parallel axiom. According as we assume the parallel axiom in the euclidean or in the hyperbolic form† we obtain the euclidean or the hyperbolic geometry.

There is also a more direct way of proving the equivalence of the obtained system with the euclidean or hyperbolic geometries; namely, the introduction of coördinates based on the group. Also it is possible to formulate the parallel-axiom in an adequate form for groups. These questions will be treated in another paper.

The subject of the present paper forms a part of the lectures which I gave in *Princeton University* during the second term of the year 1924/25. It is a great pleasure for me to express my sincerest thanks to this famous University and to the *International Education Board* whose generosity made it possible for me to hold these lectures. It is my pleasant duty also to express my indebtedness to Prof. Veblen and Prof. Alexander, whose discussions and many valuable suggestions have helped a great deal for the systematical exposition of my work.

* Hilbert, l. c., pp. 1–23, groups of axioms I–III and V.

† Hilbert, l. c., pp. 20 and 147.

UNIVERSITY OF SZEGED (HUNGARY).

June 1927.

A METHOD OF NUMERICAL SOLUTION OF THE PROBLEM OF PLATEAU.*

By Jesse Douglas.†

1. **Introduction.** We take the problem of Plateau as a boundary problem in second order partial differential equations of the elliptic type, namely, in the following formulation:

To find a function $z = f(x, y)$ which takes an assigned continuous distribution of values on a given simple contour C in the xy-plane, is continuous in the domain consisting of the region R of the plane bounded by C together with C, and which satisfies the partial differential equation of minimal surfaces,

$$(1) \qquad (1+q^2)r - 2pqs + (1+p^2)t = 0,$$

at every interior point of R.

S. Bernstein‡ and C. H. Müntz§ have demonstrated the existence of a solution of this problem under quite general assumptions as to C. The object of the present paper is to develop a method for the effective numerical computation of this solution.

We confine ourselves, in the main, to the case where the contour C is a square. However, our method may be extended to any contour of a moderate degree of regularity by covering the plane with a network of squares and approximating in the well known way to the region bounded by the contour by a region composed of squares of the network. This idea is developed in more detail at the end of § 3.

2. **Theoretical basis of solution.** Our method of numerical solution of the Plateau problem is based on the following simple theory.

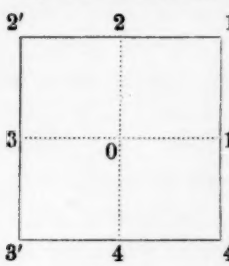


Fig. 1.

Let the figure represent a square of side $2h$ interior to the region R, the sides of the square being parallel to the coördinate axes,‖ and 1, 2, 3, 4 denoting the midpoints of the sides, 0 the center. Then in the limit as $h \to 0$, the first and second partial derivatives $p, q; r, s, t$ are equal respectively to the following difference quotients

* Received June 23, 1927; presented to the American Mathematical Society, May 7, 1927.

† National Research Fellow in Mathematics.

‡ *Sur la généralisation du problème de Dirichlet* (II), Math. Ann., vol. 69 (1910), pp. 82–136.

§ *Die Lösung des Plateauschen Problems über konvexen Bereichen*, Math. Ann., vol. 94 (1925), pp. 53–96.

‖ This implies no real limitation on the orientation of the square, since (1) is invariant under rotation of axes.

$$p = \frac{z_1 - z_3}{2h}, \quad q = \frac{z_2 - z_4}{2h}, \tag{3}$$

$$r = \frac{z_1 + z_3 - 2z_0}{h^2}, \quad s = \frac{(z_{1'} + z_{3'}) - (z_{2'} + z_{4'})}{4h^2}, \quad t = \frac{z_2 + z_4 - 2z_0}{h^2}. \tag{4}$$

The subscript of z denotes the point in the figure to which that value of z belongs.

Writing the values (3), (4) in equation (1), we have

$$\left\{1 + \frac{(z_2 - z_4)^2}{4h^2}\right\} \frac{z_1 + z_3 - 2z_0}{h^2} - \frac{1}{8} \frac{(z_1 - z_3)(z_2 - z_4)[(z_{1'} + z_{3'}) - (z_{2'} + z_{4'})]}{h^4} + \left\{1 + \frac{(z_1 - z_3)^2}{4h^2}\right\} \frac{z_2 + z_4 - 2z_0}{h^2} = 0 \tag{5}$$

as a finite difference approximation, good for small values of h, to the partial differential equation which z must satisfy. This is a linear equation for z_0 which, solved, gives

$$\begin{aligned} z_0 = \Big\{ & 4h^2(z_1 + z_2 + z_3 + z_4) + (z_1 + z_3)(z_2 - z_4)^2 + (z_2 + z_4)(z_1 - z_3)^2 \\ & - \frac{1}{2}(z_1 - z_3)(z_2 - z_4)[(z_{1'} + z_{3'}) - (z_{2'} + z_{4'})] \Big\} \\ & \div \{16h^2 + 2(z_1 - z_3)^2 + 2(z_2 - z_4)^2\}. \end{aligned} \tag{6}$$

Thus the value of z at the center of a small square is expressed in terms of its values at the vertices and midpoints of sides.

This explains the essence of the matter, but a formula more convenient or purposes of computation is the following:

$$\begin{aligned} z_0 = & \frac{a(1 + 2 + 3 + 4) + b(1' + 2' + 3' + 4')}{4(a + b)} \\ & + \frac{a + 2b}{4(a + b)} \frac{\begin{vmatrix} (1 + 3) - (2 + 4) & (1 - 3)\ (2 - 4) \\ (1' + 3') - (2' + 4') & (1' - 3')\ (2' - 4') \end{vmatrix}}{8h^2 + \begin{vmatrix} 1' - 3' & 2' - 4' \\ 1 - 3 & 2 - 4 \end{vmatrix}}. \end{aligned} \tag{7}$$

Here $1, 2, 3, 4, 1', 2', 3', 4'$ denote the values of z at the points represented by these symbols (except in $4(a+b)$ and $a+2b$ where the coefficients have their ordinary numerical significance), and a, b may be chosen as any two positive or zero numbers, but not both zero. The advantage of (7) consists in the fact that the first term, which involves only linear operations and is therefore easy to compute, gives the principal part of the value

of z_0; while the second term, whose computation is more difficult, is in the nature of a small correction and in most cases need be calculated only to one significant figure.

To verify (7), let us proceed *a posteriori*, and find an expression for the fraction with determinant numerator appearing in the right member. Using the finite form of Taylor's expansion, and regarding terms beyond h^2 as negligible, we have

$$\begin{aligned} 1 &= z_0 + ph + \frac{1}{2} rh^2, \\ 2 &= z_0 + qh + \frac{1}{2} th^2, \\ 3 &= z_0 - ph + \frac{1}{2} rh^2, \\ 4 &= z_0 - qh + \frac{1}{2} th^2; \end{aligned} \tag{8}$$

and

$$\begin{aligned} 1' &= z_0 + (p+q)h + \frac{1}{2}(r+2s+t)h^2, \\ 2' &= z_0 - (p-q)h + \frac{1}{2}(r-2s+t)h^2, \\ 3' &= z_0 - (p+q)h + \frac{1}{2}(r+2s+t)h^2, \\ 4' &= z_0 + (p-q)h + \frac{1}{2}(r-2s+t)h^2. \end{aligned} \tag{9}$$

By the substitutions (8) and (9), we find

$$\frac{\begin{vmatrix} (1+3)-(2+4) & (1-3)\ (2-4) \\ (1'+3')-(2'+4') & (1'-3')\ (2'-4') \end{vmatrix}}{8h^2 + \begin{vmatrix} 1'-3' & 2'-4' \\ 1-3 & 2-4 \end{vmatrix}} = h^2 \frac{(q^2-p^2)(r-t) - 4pqs}{2+p^2+q^2}. \tag{10}$$

Now (1) can be written

$$r + t + \frac{(q^2-p^2)(r-t) - 4pqs}{2+p^2+q^2} = 0. \tag{11}$$

It remains to find an expression for $r+t$. Add together the equations of (8) and (9) respectively:

$$1+2+3+4 = 4z_0 + (r+t)h^2, \tag{12}$$

$$1'+2'+3'+4' = 4z_0 + 2(r+t)h^2. \tag{13}$$

If we combine (12) and (13) linearly, assigning the weight $a \geqq 0$ to (12) and the weight $b \geqq 0$ to (13), avoiding the combination $a = 0$, $b = 0$, we obtain

$$(14)\quad (r+t)h^2 = \frac{a(1+2+3+4)+b(1'+2'+3'+4')}{a+2b} - \frac{4(a+b)}{a+2b} z_0.$$

Making the substitutions (14) and (10) in (11), after multiplying (11) by h^2, we get a linear equation for z_0 whose solution is (7).

It makes no difference in the ultimate result, and very little in the intermediate calculations, which values obeying the above stated conditions we assign to a and b,—a convenient choice, suggested by analogy with the Dirichlet problem, is $a = 4$, $b = 1$,* which makes (7)

$$(15)\quad z_0 = \frac{4(1+2+3+4)+(1'+2'+3'+4')}{20} + \frac{3}{10} \frac{\begin{vmatrix} (1+3)-(2+4) & (1-3)\ (2-4) \\ (1'+3')-(2'+4') & (1'-3')\ (2'-4') \end{vmatrix}}{8h^2 + \begin{vmatrix} 1'-3' & 2'-4' \\ 1-3 & 2-4 \end{vmatrix}}.$$

3. **Plan of solution.** Our method of numerical solution of the problem formulated in § 1 is based on formula (15), which enables us to calculate the value at the center of a small square of any solution of (1) by means of its values at the vertices and midpoints of sides.

First approximation. Let any square S be given and any continuous distribution of values over the points of its perimeter. Divide S into four equal squares as indicated in Fig. 1. The values of the required function are known at the midpoints and vertices 1, 2, 3, 4, 1′, 2′, 3′, 4′. By (15) we can then calculate a (first approximating) value z_0 for the center.

Second approximation. Next, let S be divided into 16 equal squares by inserting new lines of subdivision half-way between the lines already present in the first subdivision. Let the lines appearing in the previous

* The approximating formula corresponding to (7) for the Dirichlet problem is $z_0 = \{a(1+2+3+4)+b(1'+2'+3'+4')\} \div 4(a+b)$, and here the choice $a = 4$, $b = 1$ can be shown to have a definite theoretical advantage over any other. For, when $a:b = 4:1$, the difference between the two members of the above equation is of the order of h^6, whereas otherwise it is of the order of h^4.

It may be remarked that the problem of Plateau reduces to that of Dirichlet when the given contour deviates infinitely little from planarity. This depends essentially on the fact that the equation (1) of minimal surfaces, applied to the surface $z = \varepsilon q(x, y)$ deviating infinitely little from the xy-plane, becomes, after neglecting terms higher than the first degree in ε, Laplace's equation, $q_{xx} + q_{yy} = 0$.

subdivision be drawn full, the new lines broken. We obtain a figure such as Fig. 4.

We have constructed a certain net over the square S. Let us refer to the vertices of the square meshes of the net as nodes, which we may distinguish as *boundary nodes* and *interior nodes** according to an obvious definition. We shall classify the interior nodes into *closed*, *open*, and *mixed*, as follows: *closed* node, intersection of a full line of the net with a full line; *open* node intersection of a broken line with a broken line; *mixed* node, intersection of a full line with a broken line.

Values of z are assigned to the boundary nodes by hypothesis, and to the closed nodes† by the previous approximation. We begin the second approximation by finding tentative values at the mixed nodes by graphical interpolation between the values for the closed and boundary nodes.

Each interior node ν is the center of a square consisting of four meshes of our net. The nodes located at the midpoints of the sides and the vertices of this square will be said to *surround* ν. Then by the formula (15) definite values of z can be calculated for the open nodes, since each open node is surrounded solely by boundary, closed, and mixed nodes, and values of z have been assigned to all such nodes.

We next proceed to the mixed nodes, taking these in some systematic order, and calculating for each one, by means of (15), a new z-value to replace the tentative one assigned by graphical interpolation. From these we pass to the closed nodes, again using (15) to calculate a new value to replace the one found in the previous approximation. Then with these new values we go back to the open nodes, and so on, taking open, mixed, and closed nodes in cyclic repetition. Always in calculating a z-value at a node in terms of those at the surrounding nodes we use the latest value obtained for each surrounding node.

The procedure just described is illustrated in Fig. 4, where the successive z-values attached to each interior node are written in the different quadrants about the node in the usual order of the quadrants.

This stage of our problem may be formulated as follows: *Given a value of z at each boundary node of the net over the square S. To write a value z_ν on each interior node ν so that each z_ν shall be related to the surrounding*

---

* We borrow these terms from the paper of Phillips and Wiener, *Nets and the Dirichlet problem*, Journ. Math. and Phys. vol. 2 (1923), pp. 105–124, of which we became aware only after having completed the work of the present paper. The terminology of *closed*, *open*, and *mixed* nodes, introduced immediately hereafter is our own.

† There is only one of these at this stage, but there will be more in later approximations. We describe our procedure in the general terms of a passage from the nth to the $(n+1)$th approximation.

z's by formula (15). The method above described may be regarded as a technique for doing this by successive approximations.

The method results in attaching to each interior node ν a sequence of numbers $z_\nu^1, z_\nu^2, z_\nu^3, \cdots$. In ordinary cases these will converge to a definite limit z_ν, which is the value that we attach to the node ν in this second approximation.

Third approximation. This is illustrated in Fig. 5. We divide S into 64 equal squares by means of new lines of subdivision inserted half-way between the lines present in the second approximation. The old lines are drawn full, the new lines broken; as before wc have open, closed, and mixed nodes. z-values are assigned to the closed nodes by the previous approximation. We can find initial z's for the mixed nodes by graphical interpolation. Then we proceed as before from open nodes, to mixed nodes, to closed nodes, and back again cyclically. In this way we obtain a value of z for each of the 49 interior nodes, and can say that we have effected the numerial solution of 49 cubic equations in that many unknowns.

And so we can continue, using a sequence of nets each four times as fine as the preceding, until we consider the net so fine that a function defined on its nodes is practically defined throughout the square (linear interpolation). Ordinarily the net of 64 meshes will be quite sufficient for this purpose.

The spirit of this paper being entirely numerical, we do not concern ourselves with theoretical questions of convergence, which are besides too difficult for us to deal with. Whether convergence occurs or not is indicated practically in each particular case by the behavior of the numbers obtained in the course of the computation.

A complete theoretical discussion would require: first, a proof of the convergence of the process of successive approximation described above for the solution of the system of algebraic equations which results by applying (15) to each interior node; second, a discussion of the behavior of the so defined function of the nodes of the net as we make the net finer and finer, to show that there results a continuous limit function over the square S; third, a consideration of whether this limit function satisfies the partial differential equation (1), to which (15) is a finite difference approximation; and fourth, a uniqueness discussion. A disposal of these considerations would constitute, for the particular case of the square, an alternative existence proof to those of Bernstein and Müntz for the problem of Plateau.

Our procedure can be adapted without much additional difficulty to the case of a general contour C bounding a region R. We begin by covering the plane with a network of squares. Let us define *boundary*

node as one which lies on C, or which lies in R but at least one of whose surrounding nodes does not belong to $R = \bar{R} + C$. An *interior node* will be one which lies in R and all of whose surrounding nodes belong to R. To each boundary node we assign, say the z-value belonging to the nearest intersection of C with the net. We then have to solve the algebraic problem of finding values z_ν for the interior nodes so that each z_ν shall be related to the surrounding z's by (15). This can be done by a simple modification of the successive approximation process used for the square S. A rough but workable method of obtaining first approximating values is to visualize the requisite minimal surface and guess at its z-values for the nodes of an initial net with fairly large meshes. These guesses may then be corrected by application of (15). This will furnish the starting point for computation with a net four times as fine, etc.

4. **Application to a known case: the catenoid.** We check our method and also test its practical efficacy by applying it first to a known, then to an unknown case.

As known case we select the catenoid

$$z = \cosh^{-1}\sqrt{x^2+y^2}, \tag{16}$$

which is readily verified to satisfy the partial differential equation (1) of minimal surfaces.

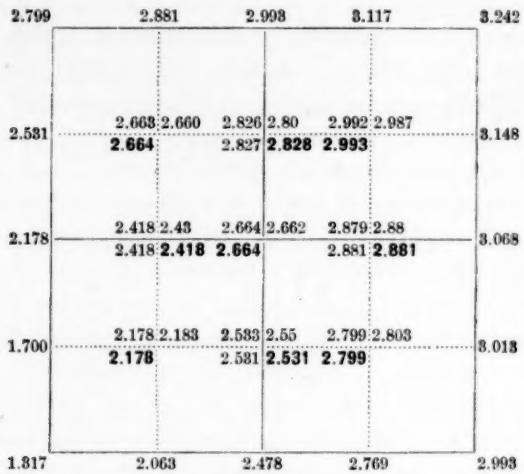


Fig. 2.

For contour C in the xy-plane we choose the square bounded by the lines $x = 2$, $x = 10$, $y = 0$, $y = 8$. We assign to the points of the perimeter of this square the values of z defined by (16). Then we apply the method described in the preceding section to compute a function $z = f(x, y)$

obeying (1) in the interior of S and attaching continuously to the assigned values on the boundary.

We carry the calculations only through the second approximation; the results are given in Fig. 2. The successive numerical approximations are written in the different quadrants around each interior node in the usual order of the quadrants, the final values being indicated in heavy type. *These values agree to the four significant figures written with those defined by* (16) *as obtained from table of hyperbolic functions.*

5. **Numerical solution of a new particular case of the problem of Plateau.** As a new case to which to apply our method, we choose the contour indicated by the lines drawn full in the following figure:

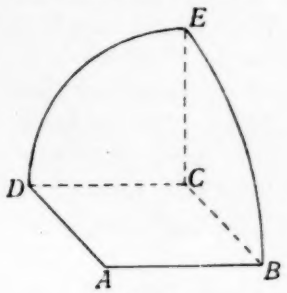


Fig. 3.

where $A\,B\,C\,D$ is a square and $B\,E$, $D\,E$ are circular quadrants with common center at C and planes perpendicular to the plane of the square.

The side of the square is taken as 20. The successive approximations (second and third) are given in Figs. 4, 5. The values in heavy type are taken as the final approximations.

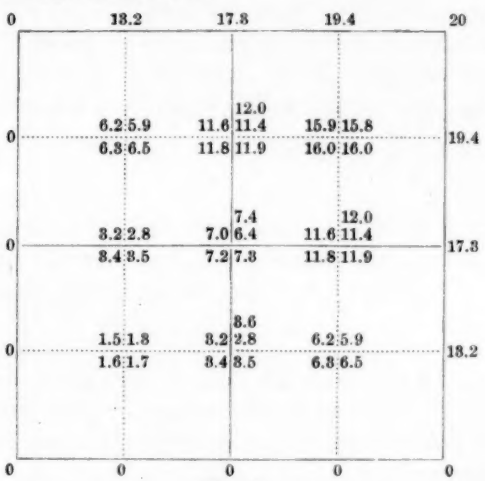


Fig. 4.

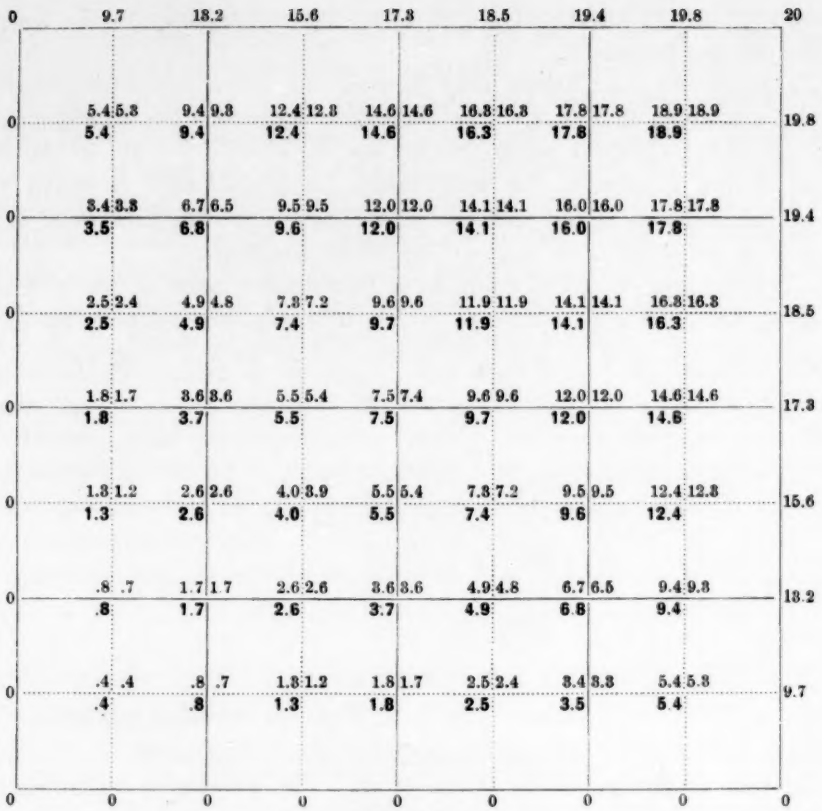


Fig. 5.

Fig. 6 is a parallel projection of the resulting minimal surface, showing the sections parallel to $D\,C\,E$ and $B\,C\,E$.

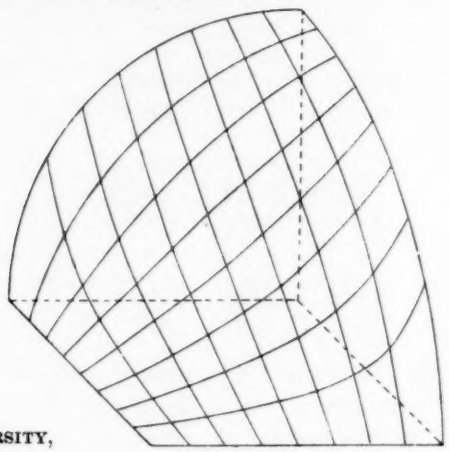


Fig. 6.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

MINIMIZING TWO TYPES OF DEFINITE INTEGRAL.*

By Paul R. Rider.

Introduction. Let us consider the two definite integrals

$$I = \int_{x_1}^{x_2} y^n (1 + y'^2)^m \, dx, \qquad J = \int_{x_1}^{x_2} y'^n (1 + y^2)^m \, dx.$$

If n is not integral we assume $y > 0$ in I and $y' > 0$ in J.

For the case $m = 1/2$ in I we have as subcases the integrals involved in the following interesting problems of the calculus of variations:† $n = -1/2$, the brachistochrone; $n = 0$, the shortest path; $n = 1/2$, the trajectory of a particle under the action of gravity; $n = 1$, the minimum surface of revolution. The integral I is thus a further generalization of an integral which itself includes a number of well-known integrals as special cases.

The integral J is of interest because certain results which are developed for it may be applied to the integral

$$K = \int_{x_1}^{x_2} p'^n (1 + p^2)^m \, dx, \qquad p = y',$$

previously considered by the writer. (See reference in § 4.) It may be remarked that for $n = -1$ in K we have as subcases the integrals arising in the following problems: $m = 1$, the curve which with its caustic encloses a minimum area; $m = 2$, the curve which with its evolute encloses a minimum area; $m = 3/2$, the curve of minimum mean radius of curvature. If, in K, $m = \frac{1}{2}(1 - 3n)$ we have the integral $\int \varkappa^n \, ds$ ($\varkappa$ = curvature) first considered by Euler.‡

It therefore appears that the integrals I and J are of sufficient interest and importance to warrant developing necessary and sufficient conditions for giving them their maximum or minimum values.

---

* Received August 29, 1927.

† See, for example, Goursat, Cours d'analyse mathématique, vol. 3, pp. 555–559; or Hadamard, Leçons sur le calcul des variations, pp. 124–131.

‡ See Ostwald's Klassiker der exakten Wissenschaften, no. 46, p. 76. See also Radon, Über das Minimum des Integrals $\int_{s_0}^{s_1} F(x, y, \theta, \varkappa) \, ds$, Sitzber. Akad. der Wiss. Wien, Math.-nat. Klasse, Bd. 119, Abt. IIa (1910), p. 61.

1. **The extremals for I and J.** The Euler-Lagrange equation for the integral I is, since its integrand, which we shall denote by f, does not contain x explicitly,*

$$f - y' f_{y'} \equiv f\left(1 - \frac{2my'}{1+y'^2}\right) = \text{const.} = a^n,$$

or

$$y^n (1+y'^2)^m \frac{1-(2m-1)y'^2}{1+y'^2} = a^n,$$

from which we get, by letting $y' = p$ and solving for y,

$$y = a(1+p^2)^{(1-m)/n}[1-(2m-1)p^2]^{-1/n}. \tag{1}$$

Since $p = dy/dx$, we find

$$x = \int \frac{1}{p}\frac{dy}{dp}\,dp + b, \tag{2}$$

and (1) and (2) are parametric equations of the extremals for the integral I.

For the integral J, whose integrand g also does not contain x explicitly, the Euler-Lagrange equation is

$$g - y' g_{y'} \equiv g(1-n) = \text{const.} = a^n,$$

or

$$y' = (1-n)^{-1/n} a (1+y^2)^{-m/n},$$

from which we get

$$x = \frac{1}{a}(1-n)^{1/n}\int (1+y^2)^{m/n}\,dy + b$$
$$= a_1 \int (1+y^2)^{m/n}\,dy + b,$$

where a_1 involves n. This is the equation of the extremals for J.

2. **Sufficient conditions for minimizing the integral I.** In order to develop sufficient conditions for minimizing† the integral I, let us consider the difference between the value of the integral for a comparison curve $y = y(x) + \eta(x)$ and its value for an extremal. Denoting this difference by ΔI we find, by using Taylor's theorem, that

$$\Delta I = \int_{x_1}^{x_2} f\left(\frac{n}{y}\eta + \frac{2my'}{1+y'^2}\eta'\right)dx$$
$$+ \frac{1}{2}\int_{x_1}^{x_2} \bar{f}\left\{\frac{n(n-1)}{\bar{y}^2}\eta^2 + \frac{4mn\bar{y}'}{\bar{y}(1+\bar{y}'^2)}\eta\eta' + \frac{2m[1+(2m-1)\bar{y}'^2]}{(1+\bar{y}'^2)^2}\eta'^2\right\}dx,$$

* See Bolza, Vorlesungen über Variationsrechnung, p. 32.

† Throughout this paper sufficient conditions refer to weak variations only.

in which $\bar{y} = y + \theta\eta$, $\bar{y}' = y' + \theta\eta'$ $(0 < \theta < 1)$, $\bar{f} = f(\bar{y}, \bar{y}')$. It can readily be shown that the first integral vanishes, since y and y' refer to the extremal. The integrand of the second integral may be written in the form $\bar{f}Q$,

$$Q \equiv n(n-1)\alpha^2 + 4mn\bar{y}'\alpha\beta + 2m[1 + (2m-1)\bar{y}'^2]\beta^2,$$

where $\alpha = \eta/\bar{y}$, $\beta = \eta'/(1+\bar{y}'^2)$.

It follows that ΔI will be positive or negative according as $\bar{f}Q$ is greater than or less than zero. If we assume $\bar{f} > 0$, then a sufficient condition that I be minimized by the extremal is that the quadratic form Q be positive definite;* similarly, a sufficient condition that I be maximized is that Q be negative definite. Q will be definite if its discriminant is negative, i. e. if

$$4m^2n^2\bar{y}'^2 - 2m[1 + (2m-1)\bar{y}'^2]n(n-1) < 0.$$

This condition can be reduced to the form

$$mn[1 - n + (2m+n-1)\bar{y}'^2] < 0. \tag{3}$$

We shall consider four cases.

Case I. $m > 0$, $n < 0$.

The inequality (3) holds for all $\bar{y}'$ if $2m+n-1 \geqq 0$; when $2m+n-1 < 0$, (3) holds if and only if

$$\bar{y}'^2 < (n-1)/(2m+n-1).$$

In Case I the coefficient of α^2 in the quadratic form Q is positive, hence the extremals will always *minimize* the integral I if Q is definite.

Case II. $m > 0$, $n > 0$.

The inequality (3) holds if and only if $\bar{y}'^2 \lessgtr (n-1)/(2m+n-1)$ according as $2m+n-1 \gtrless 0$.

The coefficient $n(n-1)$ of α^2 in Q will be negative if $n < 1$ and positive if $n > 1$. But $n > 1$ implies $2m+n-1 > 0$. Therefore

If $2m+n-1 < 0$, a sufficient condition for a *maximum* is $\bar{y}'^2 > (n-1)/(2m+n-1)$.

If $2m+n-1 > 0$, a necessary and sufficient condition for an *extremum* is $\bar{y}'^2 < (n-1)/(2m+n-1)$; the extremum will be a *minimum* since, if n were less than 1, $\bar{y}'^2$ would have to be negative, which is impossible.

Case III. $m < 0$, $n > 0$.

* See Bôcher, Introduction to higher algebra, p. 150.

The inequality (3) holds if and only if $\bar{y}'^2 \gtrless (n-1)/(2m+n-1)$ according as $2m+n-1 \gtrless 0$.

The coefficient $n(n-1)$ of α^2 in Q will be negative if $n<1$ and positive if $n>1$. But $n<1$ implies $2m+n-1<0$. Therefore

If $2m+n-1>0$, a sufficient condition for a *minimum* is $\bar{y}'^2>(n-1)/(2m+n-1)$.

If $2m+n-1<0$, a sufficient condition for an *extremum* is $\bar{y}'^2<(n-1)/(2m+n-1)$; the extremum will be a *maximum*, since $n>1$ implies $\bar{y}'^2<0$.

Case IV. $m<0$, $n<0$.

This case implies $2m+n-1<0$.

The inequality (3) holds if and only if $\bar{y}'^2>(n-1)/(2m+n-1)$.

Since $n(n-1)>0$, the quadratic form Q, if definite, will be positive definite, and the extremals will *minimize* I.

Thus it is seen that if we consider only curves whose slopes satisfy the proper inequality we are able to state sufficient conditions that the integral I be minimized of maximized.

The cases $m=0$ and $n=0$ are not of much interest.

3. **Sufficient conditions for minimizing the integral** J. The difference between the value of J taken over a comparison curve and its value taken over an extremal can be shown to be equal to

$$\Delta J = \frac{1}{2}\int_{x_1}^{x_2} \bar{g}\,\{n(n-1)\alpha^2+4m\bar{y}\,\alpha\beta+2m[1+(2m-1)\bar{y}^2]\,\beta^2\}\,dx,$$

where here $\alpha=\eta'/\bar{y}'$, $\beta=\eta/(1+\bar{y}^2)$, and $\bar{g}=g(\bar{y},\bar{y}')$.

Assuming $\bar{g}>0$, we see that the quadratic form in the $\{\ \}$ will be definite if

$$mn[1-n+(2m+n-1)\bar{y}^2]<0.$$

The discussion then follows precisely that for the integral I if we replace $\bar{y}'$ by $\bar{y}$.

4. **Application to particular cases.** In an earlier paper* the writer has derived the equations of the extremals for the integral

$$K=\int_{x_1}^{x_2} p'^m\,(1+p^2)^m\,dx, \qquad p=y'.$$

The sufficient conditions derived for the integral J apply directly here if we write $\bar{p}$ in place of $\bar{y}$. That is, they are exactly the same as the sufficient conditions for the integral I.

* Rider, The minimum area between a curve and its caustic. Bull. Amer. Math. Soc., vol. 27, (1921), pp. 279–284, § 3, A more general problem, p. 283.

A special case of K previously treated in some detail* is that in which $n = -1$. In this case the inequality (3) reduces to

$$m[1 + (m-1)\bar{y}'^2] > 0.$$

From this we are able slightly to extend some of the results of the paper just referred to. Thus, in the statements in the first two paragraphs at the top of p. 171, the relations $|p| \leqq 1$ and $|p| \geqq 1$ can be replaced by the relations $|\bar{p}| < 1/(1-m)$ and $|\bar{p}| > 1/(1-m)$ respectively, which are more inclusive, since in the first case $m > 0$ and in the second case $m < 0$.

Finally, let us consider Euler's problem† of minimizing the integral $\int x^n \, ds$. This, as already stated, is the special form of the integral K in which $m = \frac{1}{2}(1-3n)$.

If $n < 0$, then $m > 0$ and we have Case I. The quantity $2m + n - 1 = -2n > 0$ and the sufficient condition that the integral be *minimized* is satisfied.

If $0 < n < 1/3$, then $m > 0$ and we have Case II. Here $2m + n - 1 = -2n < 0$, $n(n-1) < 0$, and the sufficient condition for a *maximum* is $\bar{y}'^2 > (1-n)/2n$.

If $n > 1/3$, then $m < 0$ and we have Case III. The quantity $2m + n - 1 = -2n < 0$. The condition $\bar{y}'^2 < (1-n)/2n$ can not be fulfilled if $n \geqq 1$. If $n < 1$, the condition may be fulfilled. If it is, the integral will be *maximized*.

---

* Rider, On the minimizing of a class of definite integrals, Annals of Math. (2), vol. 24, (1922), pp. 167-174.

† See Ostwald's Klassiker, no. 46, loc. cit.; also Radon, loc. cit.

WASHINGTON UNIVERSITY ST. LOUIS, MO.

ON TWO TYPES OF PLANE RATIONAL CURVE.*

By Harold Hilton.

1. The types considered are those traced out by the point (x, y, z), where

$$x : y : z = \sin(n+p)\varphi : \sin n\varphi : \sin(n-p)\varphi \tag{1}$$

or

$$x : y : z = \cos(n+p)\varphi : \cos n\varphi : \cos(n-p)\varphi . \tag{2}$$

Here n and p are positive integers prime to one another with $n > p$. In much of what follows we may suppose $p > n$, but we shall confine ourselves to the case $n > p$.

To save space the properties of the curve are merely put on record without detailed proof.

The curves are both rational (unicursal), for $x : y : z$ are expressible rationally in terms of the parameter $2\cos\varphi$.

It will be convenient to use also another set of homogeneous coördinates (ξ, η, ζ) and a Cartesian set (X, Y), connected with (x, y, z) by

$$\begin{aligned} X &= y/(x+z), \qquad Y = z/(x+z), \\ \xi : \eta : \zeta &= (x+2y+z) : (x-2y+z) : (x-z) \\ &= (1+2X) \quad : (1-2X) \quad : (1-2Y). \end{aligned}$$

The points $A, B, C, D, E, F, I, J, U, X, W$ referred to later are those with the ξ, η, ζ coördinates†

$$\begin{aligned} &A(1, -1, 0), \quad B(0, 0, 1), \quad C(1, 1, 1), \quad D(1, 1, -1), \\ &E(1, 0, 0), \quad F(0, 1, 0), \quad I(2n, 0, p), \quad J(0, 2n, p), \\ &U(p, p, n), \quad X(2n, 2n, p), \quad W(4np, 4np, 4n^2+p^2). \end{aligned}$$

The conic

$$y^2 = zx, \text{ i. e. } X^2 + Y^2 = Y \text{ or } \zeta^2 = \xi\eta \tag{3}$$

meets the curves only at C, E, F.

The intersections of the curves with AD, BC, AC, AB are all real.

The first polar curve of B is obtained by replacing p in (1) and (2) by $p-1$ if p is odd, and by $p-2$ if p is even.

* Received September 9, 1927.

† Proc. Lond. Math. Soc., (2), vol. 20, p. 93. The notation $A, B, \cdots, W$ is derived from this paper.

The birational transformation

$$X' = X, \quad Y' = 1 - X^2/Y$$

converts each of the curves (1) and (2) into the same curve with $n - p$ instead of n.

The birational transformation

$$\xi' = 1/\eta, \quad \eta' = 1/\xi, \quad \zeta' = 1/\zeta$$

converts each of the curves (1) and (2) into the other.

These transformations may be used to obtain new properties of the curves from those given in the following sections.

2. Consider first the curve (1).

Included among the type is the curve of degree n ("n-ic") with a multiple point of order $n-1$ ("$(n-1)$-ple point") and a conic of $2n$-point contact, for which $p = 1$.

Also the k-ic with a $(k-1)$-ple point O meeting a conic at $(2k-1)$ consecutive points and at a point of contact of a tangent from O to the conic is given by $n = 2k-1$, $p = 2$.

Any conic can be put into the form (1) either with $n = 2$, $p = 1$ or $n = 3$, $p = 2$.

Any crunodal cubic can be put into the form with either with $n = 3$, $p = 1$ or $n = 5$, $p = 2$.

Any acnodal cubic can be projected into the form

$$(a - r)/(a + r) = -\tan^6 \frac{1}{2}\theta$$

derived from the case $n = 3$, $p = 1$ (see § 3 (7)).

3. Suppose first that p is *odd*.

The degree is $(n + p - 1)$, the class is $(2n + p - 3)$.

At B is an $(n-1)$-ple point with real distinct tangents

$$\xi \sin^2(k\pi/2n) + \eta \cos^2(k\pi/2n) = 0, \qquad k = 1, 2, \cdots, n-1.$$

Each of E and F is a $\frac{1}{2}(p-1)$-ple point at which the curve has $\frac{1}{2}(p-1)$ real linear branches all touching the conic (3).

At C is a p-ple point each of whose branches touches the conic (3). It is a higher singularity equivalent to $\frac{1}{2}(p-1)(2n-1)$ double points, of which $(p-1)$ count as cusps in reckoning the class.

The equation of the curve is

$$\left(\frac{\zeta - \sqrt{(\xi\eta)}}{\zeta + \sqrt{(\xi\eta)}}\right)^p = \left(\frac{\sqrt{\xi} - \sqrt{\eta}}{\sqrt{\xi} - \sqrt{\eta}}\right)^{2n} \tag{4}$$

or

$$Y^p u_n - {}^pC_1\, Y^{p-1} X^2 u_{n-1} + {}^pC_2\, Y^{p-2} X^4 u_{n-2} - \cdots = 0, \tag{5}$$

where

$$u_r \equiv 1 - a_1 X^2 + a_2 X^4 - a_3 X^6 + \cdots,$$

$a_1, a_2, a_3, \cdots$ being the positive integers such that, if $t \equiv 2\cos\theta$,

$$\sin r\theta \operatorname{cosec}\theta \equiv t^{r-1} - a_1 t^{r-3} + a_2 t^{r-5} - \cdots. \tag{6}$$

Writing

$$\zeta = 0, \qquad \xi = re^{\theta i}, \qquad \eta = re^{-\theta i},$$

we see that the curve is projectable (by an unreal projection) into the curve with polar equation

$$\{(a-r)/(a+r)\}^p = (-1)^n \tan^{2n}\frac{1}{2}\theta, \tag{7}$$

where B is now the pole, C the point $(a, 0)$, E and F the circular points at infinity.

The equation (5) may be put into the form

$$Y = \frac{X^2}{1} - \frac{X^2}{1} - \cdots - \frac{X^2}{1} - \frac{X^2}{Y} - \frac{X^2+Y^2-Y}{Y} - \cdots \\ \cdots - \frac{X^2}{Y} - \frac{X^2+Y^2-Y}{Y}, \tag{8}$$

with $n-p$ "$\dfrac{X^2}{1}$" and $p-1$ "$\dfrac{X^2}{Y} - \dfrac{X^2+Y^2-Y}{Y}$" in the continued fraction.

4. Now let p be *even* in (1). Then n must be odd.

The degree is $\frac{1}{2}(n+p-1)$, the class is $\frac{1}{2}(2n+p-4)$. At B is a $\frac{1}{2}(n-1)$-ple point with real distinct tangents.

If $\frac{1}{2}p$ is odd,

$\frac{1}{4}(p-2)$ real linear branches touch the conic (3) at E, while $\frac{1}{4}(p-2)$ real linear branches touch the conic (3) at F and another linear branch passes through F.

If $\frac{1}{2}p$ is even,

$\frac{1}{4}(p-4)$ real linear branches touch the conic (3) at E and another linear branch passes through E, while $\frac{1}{4}p$ real linear branches touch the conic at F.*

At C is a $\frac{1}{2}p$-ple point, each of whose branches touches the conic (3). It is equivalent to $\frac{1}{4}(p-2)(n-1)$ double points, of which $\frac{1}{2}(p-2)$ count as cusps in reckoning the class.

* Of course E and F may be interchanged in these statements.

The equation of the curve is got by equating the square root of each side of (4); or on replacing in (5) p by $\frac{1}{2}p$ and $u_n, u_{n-1}, u_{n-2}, \cdots$ by $u_n, u_{n-2}, u_{n-4}, \cdots$, where u_r is now

$$(1+2X)^{\frac{1}{2}(r-1)} - a_1(1+2X)^{\frac{1}{2}(r-3)}X + a_2(1+2X)^{\frac{1}{2}(r-5)}X^2 - \cdots.$$

In (8) we replace Y on the left by $Y+X$, and replace the first $\frac{X^2}{1}$ by $\frac{X(1+2X)}{1+X}$, having then $\frac{1}{2}(n-p-1)$ "$\frac{X^2}{1}$" and $\frac{1}{2}(p-2)$ "$\frac{X^2}{Y} - \frac{X^2+Y^2-Y}{Y}$".

5. Let now n and p be any constant quantities. Writing

$$(9) \qquad \alpha \equiv \sin(n+p)\varphi \operatorname{cosec} n\varphi, \qquad \beta \equiv \sin(n-p)\varphi \operatorname{cosec} n\varphi,$$

so that

$$(10) \qquad \alpha = x/y, \qquad \beta = z/y,$$

the tangent at any point of (1) is

$$\lambda x + \mu y + \nu z = 0,$$

where

$$(11) \qquad \begin{aligned} \lambda : \mu : \nu = {} & \{-(2n-p)\alpha\beta + p\beta^2 + 2(n-p)\} \\ & : 2\{n(\alpha+\beta)\alpha\beta - (n-p)\alpha - (n+p)\beta\} \\ & : \{-(2n+p)\alpha\beta - p\alpha^2 + 2(n+p)\}. \end{aligned}$$

The inflexions are given by

$$(12) \quad n\{(2n-p)\alpha - (2n+p)\beta\}\alpha\beta = (n-p)(2n+p)\alpha - (n+p)(2n-p)\beta.$$

From (10) and (12) we see that the inflexions all lie on a cubic

$$n\{(2n-p)x - (2n+p)z\}xz = \{(n-p)(2n+p)x - (n+p)(2n-p)z\}y^2$$

cutting the conic (3) at E, F and touching it at C, D.

It touches the curve at I and J.

From (11) and (12) we see that the inflexional tangents all touch a class-cubic

$$16n\{(n+p)\lambda - (n-p)\nu\}\lambda\nu = \{(2n+p)^2\lambda - (2n-p)^2\nu\}\mu^2.$$

This class-cubic touches the curve at I and J, passes through E and F, and touches BE and BF.

From (10) and (11) we see that the points of contact from any given point to the curve lie on a cubic through E, F, I, J.

If the point lies on BC, they lie on a conic. The points of contact of tangents from U lie on EF and IJ, and from W lie on XI and XJ.

Similarly from (10) and (11) tangents to the curve (1) at its intersection with any line all touch the same rational class-cubic (which degenerates for special positions of the line) having the given line as bitangent.

Also the tangents to (1) at its intersections with a conic touching AB and AC at B and C (other than tangents at B and C) touch a class-quartic. But if the conic goes through I and J, these tangents (other than tangents at B, C, I, J) all touch a conic.

6. The curve (2) is included in (1) if p is even, as is seen by putting $\varphi = \frac{1}{2}\pi + \psi$.

Suppose p is *odd*.

The degree is $(n+p)$ and the class is $(2n+p-1)$.

At B is an n-ple point with real distinct tangents.

Each of E and F is a $\frac{1}{2}(p+1)$-ple point at which $\frac{1}{2}(p-1)$ linear branches touch the conic (3) and one linear branch cuts the conic.

At C is a p-ple point each of whose branches touches (3). It is a higher singularity equivalent to $\frac{1}{2}(p-1)(2n-1)$ double points, of which $p-1$ count as cusps in reckoning the class.

The type includes the $(n+1)$-ic with an n-ple point O meeting a conic at $2n$ consecutive points and at the points of contact of the tangents from O to the conic. For this curve $p = 1$.

Any crunodal cubic is included in the type. It is given by $n = 2$, $p = 1$.

The discussion of type (2) is otherwise much the same as that of type (1) with the following changes.

In equation (4) insert a minus sign after the "$=$". Replace (6) by

$$2\cos r\theta = t^r - a_1 t^{r-2} + a_2 t^{r-4} - \ldots.$$

Change $(-1)^n$ into $(-1)^{n+1}$ in (7).

Replace (9) by

$$\alpha \equiv \cos(n+p)\varphi \sec n\varphi, \qquad \beta \equiv \cos(n-p)\varphi \sec n\varphi.$$

A CORRESPONDENCE BETWEEN MATRICES AND QUADRATIC IDEALS.*

By C. C. MacDuffee.

1. **Introduction.** Poincaré† first noted that all numbers not nilfactors of a linear associative algebra with a principal unit correspond to the elements of a certain linear group. In the present paper this correspondence of Poincaré is applied to the quadratic field to establish a p-to-1 correspondence between certain matrices and all principal ideals, p being the number (finite or infinite) of units in the field. It is shown that this same p-to-1 correspondence can be extended to hold between certain matrices and all ideals of the field.‡ In fact, a necessary and sufficient condition in order that a matrix G with rational integral elements correspond to an ideal is that the *class matrix* $A = G\Gamma G^{-1}$ have rational integral elements, Γ being a matrix fixed for the field. Two class matrices A_1 and A_2 are *similar* if and only if there is a matrix C with rational integral elements and of determinant ± 1 such that $A_1 = CA_2C^{-1}$. All the ideal matrices G_i of a class determine similar class matrices, and conversely, so that the number of dissimilar class matrices A_i is the class number of the field.

If we choose one class matrix A to represent each ideal class, we find that in every class there are two basic matrices B_1 and B_2 such that all ideals of the class correspond to the matrices $xB_1 + yB_2$, where x and y range independently over all rational integers, and conversely. It is possible to choose A in such a way that B_1 corresponds to the canonical basis of a reduced ideal, using the usual definitions of these terms.

This method of approach to the theory of ideals through matrices seems to be of considerable power in all connections in which the product of ideals is not involved. It is well adapted to the calculation of ideals and ideal classes.

---

* Received September 9, 1927; presented to the American Mathematical Society, April 15, 1927.

† Paris *Comptes Rendus*, vol. 99 (1884), p. 740.

‡ Stated by Poincaré who, however, did not pursue the idea. *Jour. de l'Éc. Pol.*, 28 (1880), p. 239 and *Bull. de la Soc. Math. de France*, 13 (1885), p. 167.

2. **The correspondence of Poincaré.** The integral numbers of the quadratic field $F(\sqrt{m})$* are given by $x+y\theta$ where x and y range independently over all rational integers, and where $\theta^2 = m$ in Case I, i. e., when $m \equiv 2, 3 \pmod 4$ but $\theta^2 = \theta + m'$, $4m' = m-1$ in Case II, i. e., when $m \equiv 1 \pmod 4$.

Let us define for the respective cases

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \Gamma = \begin{pmatrix} 0 & 1 \\ m & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ m' & 1 \end{pmatrix}.$$

It is readily verified that the matrices I and Γ behave under addition, multiplication and scalar multiplication precisely as do the basal numbers 1 and θ. It follows then that the integral numbers $x+y\theta$ of the field are isomorphic with the set of matrices $xI+y\Gamma$, x and y being rational integers. We shall say that the number $a = a_1 + a_2\theta$ *corresponds* to the matrix $M(a) = a_1 I + a_2 \Gamma$ under a *Poincaré correspondence.*

3. **A correspondence between matrices and ideals.** A set of integral numbers not all zero was called an ideal by Dedekind if the set was closed under addition and subtraction, and under multiplication by all the integral numbers of the field. It is well known that every ideal has a basis,† i. e., that there exist in the ideal two numbers ω_1 and ω_2 such that the totality of numbers of the ideal is given by $x\omega_1 + y\omega_2$ where x and y range independently over all rational integers.

It is not true conversely that every pair of integral numbers ω_1, ω_2 constitute a basis for an ideal. We shall deduce necessary and sufficient conditions that they do.

Let us assume that

$$\omega_1 = g_{11} + g_{12}\theta, \qquad \omega_2 = g_{21} + g_{22}\theta$$

constitute a basis for an ideal. Evidently the g_{ij} are rational integers. Every number k of the ideal is of the form $k = k_1\omega_1 + k_2\omega_2$, while every integral number s of F is of the form $s = s_1 + s_2\theta$. Since every sk is in the ideal, there must exist rational integers l_1 and l_2 such that $sk = l_1\omega_1 + l_2\omega_2$. But in Case I

$$\begin{aligned} sk = {} & s_1k_1g_{11} + s_1k_2g_{21} + s_2k_1g_{12}m + s_2k_2g_{22}m \\ & + (s_1k_1g_{12} + s_2k_1g_{11} + s_1k_2g_{22} + s_2k_2g_{21})\,\theta \end{aligned}$$

---

* See, for instance, J. Sommer, *Vorlesungen über Zahlentheorie*, Leipzig, 1907, or the translation by A. Lévy, *Introduction à la Théorie des Nombres Algebriques*, Paris, 1911, p. 21.

† See, for instance, E. Landau, *Theorie der algebraischen Zahlen und der Ideale*, Leipzig, 1918, p. 29, Theorem 99.

while

$$l_1\omega_1 + l_2\omega_2 = l_1 g_{11} + l_2 g_{21} + (l_1 g_{12} + l_2 g_{22})\,\theta.$$

On account of the linear independence of 1 and θ, we must have

$$\begin{aligned} s_1 k_1 g_{11} + s_1 k_2 g_{21} + s_2 k_1 g_{12} m + s_2 k_2 g_{22} m &= l_1 g_{11} + l_2 g_{21}, \\ s_1 k_1 g_{12} + s_1 k_2 g_{22} + s_2 k_1 g_{11} \quad + s_2 k_2 g_{21} \quad &= l_1 g_{12} + l_2 g_{22}. \end{aligned}$$

In particular, there must exist integers l_{11}, l_{12} when $(s_1, s_2, k_1, k_2) = (0, 1, 1, 0)$ and integers l_{21}, l_{22} when $(s_1, s_2, k_1, k_2) = (0, 1, 0, 1)$. This gives the necessary conditions:

$$\begin{aligned} g_{12} m &= l_{11} g_{11} + l_{12} g_{21}, \quad & g_{11} &= l_{11} g_{12} + l_{12} g_{22}, \\ g_{22} m &= l_{21} g_{11} + l_{22} g_{21}, \quad & g_{21} &= l_{21} g_{12} + l_{22} g_{22}, \end{aligned}$$

which can be written in the form

$$(1) \qquad \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ m & 0 \end{pmatrix} = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}.$$

In Case II we have, similarly,

$$\begin{aligned} sk = {} & s_1 k_1 g_{11} + s_1 k_2 g_{21} + s_2 k_1 g_{12} m' + s_2 k_2 g_{22} m' \\ & + (s_1 k_1 g_{12} + s_2 k_1 g_{11} + s_2 k_1 g_{12} + s_1 k_2 g_{22} + s_2 k_2 g_{21} + s_2 k_2 g_{22})\,\theta \end{aligned}$$

while $l_1\omega_1 + l_2\omega_2$ has the same value as before. Consequently we must have

$$\begin{aligned} s_1 k_1 g_{11} + s_1 k_2 g_{21} + s_2 k_1 g_{12} m' + s_2 k_2 g_{22} m' &= l_1 g_{11} + l_2 g_{21}, \\ s_1 k_1 g_{12} + s_2 k_1 g_{11} + s_2 k_1 g_{12} + s_1 k_2 g_{22} + s_2 k_2 g_{21} + s_2 k_2 g_{22} &= l_1 g_{12} + l_2 g_{22}. \end{aligned}$$

Setting $(s_1, s_2, k_1, k_2) = (0, 1, 1, 0), (0, 1, 0, 1)$ in turn, we obtain the necessary conditions

$$\begin{aligned} g_{12} m' &= l_{11} g_{11} + l_{12} g_{21}, \quad & g_{11} + g_{12} &= l_{11} g_{12} + l_{12} g_{22}, \\ g_{22} m' &= l_{21} g_{11} + l_{22} g_{21}, \quad & g_{21} + g_{22} &= l_{21} g_{12} + l_{22} g_{22}. \end{aligned}$$

These can be written in matric form:

$$(1') \qquad \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ m' & 1 \end{pmatrix} = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}.$$

These conditions are in fact sufficient that ω_1, ω_2 constitute a basis for an ideal. For, let $l_{11}, l_{12}, l_{21}, l_{22}$ and $g_{11}, g_{12}, g_{21}, g_{22}$ be rational integers satisfying (1) (or (1′)). Set

$$\omega_1 = g_{11} + g_{12}\theta, \quad \omega_2 = g_{21} + g_{22}\theta.$$

The set of numbers $k = k_1\omega_1 + k_2\omega_2$ where k_1 and k_2 vary over all rational integers is evidently closed under addition and subtraction. We show that the set constitutes an ideal by showing that it is closed under multiplication by all integral numbers $s = s_1 + s_2\theta$ of F. In Case I,

$$sk = s_1k_1(g_{11} + g_{12}\theta) + s_1k_2(g_{21} + g_{22}\theta) + s_2k_1(g_{12}m + g_{11}\theta) + s_2k_2(g_{22}m + g_{21}\theta).$$

In view of (1) we may write

$$sk = (s_1k_1 + s_2k_1l_{11} + s_2k_2l_{21})\omega_1 + (s_1k_2 + s_2k_1l_{12} + s_2k_2l_{22})\omega_2$$

which is therefore in the set.

In Case II we have, similarly,

$$sk = s_1k_1(g_{11} + g_{12}\theta) + s_1k_2(g_{21} + g_{22}\theta) + s_2k_1(g_{12}m' + g_{11}\theta + g_{12}\theta) + s_2k_2(g_{22}m' + g_{21}\theta + g_{22}\theta)$$

which, in view of (1′), may be written

$$sk = (s_1k_1 + s_2k_1l_{11} + s_2k_2l_{21})\omega_1 + (s_1k_2 + s_2k_1l_{12} + s_2k_2l_{22})\omega_2.$$

We shall say that the matrix

$$G = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

corresponds to the basis $\omega_1 = g_{11} + g_{12}\theta$, $\omega_2 = g_{21} + g_{22}\theta$ of the ideal (ω_1, ω_2). Where no confusion results, we may say that G corresponds to the ideal itself.

We have now proved our fundamental theorem:

THEOREM 1. *A necessary and sufficient condition in order that* $\omega_1 = g_{11} + g_{12}\theta$, $\omega_2 = g_{21} + g_{22}\theta$ *constitute a basis for an ideal is that there exist a matrix* A *with rational integral elements such that*

$$G\Gamma = AG \tag{2}$$

where G *is the matrix corresponding to* (ω_1, ω_2), *i. e., that* G *transform* Γ *into a matrix* A *with rational integral elements.*

We note that $|G|$ is the norm of the ideal, and hence is not zero.

4. Change of basis. It is well known that the basis of an ideal is not unique, but that the totality of bases (ω_1', ω_2') is given in terms of any one basis (ω_1, ω_2) by

$$\omega_1' = c_{11}\omega_1 + c_{12}\omega_2, \qquad \omega_2' = c_{21}\omega_1 + c_{22}\omega_2 \tag{3}$$

where the c's are rational integers of determinant ± 1. If we set

$$\omega_1' = g_{11}' + g_{12}'\theta, \qquad \omega_2' = g_{21}' + g_{22}'\theta$$

and denote the matrices (g_{ij}') by G' and (c_{ij}) by C, we have under (3), $G' = CG$.

The effect of a change of basis upon (2) is seen by multiplying (2) on the left by C:

$$CG\Gamma = CAC^{-1}CG, \quad \text{or} \quad G'\Gamma = CAC^{-1}G'.$$

We shall say that two matrices A_1 and A_2 are *similar* if there exists a matrix C with rational integral elements and of determinant ± 1 such that $A_1 = CA_2C^{-1}$. Evidently the relation is reciprocal. We now have

THEOREM 2. *The effect of the change of basis* (3) *of an ideal is to replace the corresponding matrix G by CG and to transform the matrix A into the similar matrix CAC^{-1}.*

5. Principal ideals. A principal ideal $\{d\}$ is composed of all integral multiples of a fixed integral number $d = d_1 + d_2\theta$. That is, $\{d\}$ is composed of all numbers sd where $s = s_1 + s_2\theta$ and s_1 and s_2 are rational integral variables. We have for Case I

$$sd = s_1d_1 + s_2d_2m + (s_1d_2 + s_2d_1)\theta = s_1\omega_1 + s_2\omega_2$$

where

$$\omega_1 = d_1 + d_2\theta, \qquad \omega_2 = d_2m + d_1\theta.$$

Evidently

$$G = \begin{pmatrix} d_1 & d_2 \\ d_2m & d_1 \end{pmatrix} = d_1I + d_2\Gamma = M(d).$$

Thus $M(d)$ corresponds to the principal ideal $\{d\}$.

Likewise in Case II,

$$sd = s_1d_1 + s_2d_2m' + (s_1d_2 + s_2d_1 + s_2d_2)\theta = s_1\omega_1 + s_2\omega_2$$

where

$$\omega_1 = d_1 + d_2\theta, \qquad \omega_2 = d_2m' + (d_1 + d_2)\theta.$$

Evidently

$$G = \begin{pmatrix} d_1 & d_2 \\ d_2m' & d_1 + d_2 \end{pmatrix} = d_1I + d_2\Gamma = M(d).$$

Hence in this case also, $M(d)$ is a matrix corresponding to the principal ideal $\{d\}$.

Since multiplication in F is commutative, and since $M(d)$ is isomorphic with d and Γ with θ, we have

$$M(d)\Gamma = \Gamma M(d). \tag{4}$$

We have therefore proved

THEOREM 3. *The matrix $M(d)$ corresponds to a basis of the principal ideal $\{d\}$, the corresponding value of the matrix A of Theorem 2 being Γ.*

By Theorem 2 every basis of the principal ideal $\{d\}$ has a corresponding matrix $G = CM(d)$ where $|C| = \pm 1$, and therefore the matrix $A = C\Gamma C^{-1}$ is similar to Γ.

Let us determine the most general matrix commutative with Γ. We have in Case I

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ m & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ m & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

or

$$\begin{pmatrix} bm & a \\ dm & c \end{pmatrix} = \begin{pmatrix} c & d \\ am & bm \end{pmatrix}.$$

Thus $d = a$, $c = bm$, and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ bm & a \end{pmatrix} = aI + b\Gamma.$$

Likewise in Case II,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ m' & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ m' & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

$$\begin{pmatrix} bm' & a+b \\ dm' & c+d \end{pmatrix} = \begin{pmatrix} c & d \\ am'+c & bm'+d \end{pmatrix}$$

so that $c = bm'$ and $d = a+b$. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ bm' & a+b \end{pmatrix} = aI + b\Gamma.$$

Suppose now that G is an integral matrix such that

$$G\Gamma = AG, \qquad A = C^{-1}\Gamma C, \qquad |C| = \pm 1.$$

Then $CG\Gamma = \Gamma CG$ so that CG is commutative with Γ, and so is of the form $aI + b\Gamma$. Then CG, and hence G also, determines the principal ideal $\{a + b\theta\}$. Thus we have

THEOREM 4. *A necessary and sufficient condition in order that the matrix G correspond to a principal ideal is that $G\Gamma = AG$ where A is a matrix similar to Γ.*

6. Ideal classes. Two ideals (ω_1, ω_2) and (ω_1', ω_2') are said to be *equivalent* or to belong to the same *class* if there exist integral numbers k and k' of F such that $(k\omega_1, k\omega_2) = (k'\omega_1', k'\omega_2')$.

Let us set

$$\omega_1 = g_{11} + g_{12}\theta, \qquad \omega_2 = g_{21} + g_{22}\theta, \qquad k = k_1 + k_2\theta.$$

Then in Case I,

$$k\omega_1 = k_1 g_{11} + k_2 g_{12} m + (k_1 g_{12} + k_2 g_{11})\theta,$$
$$k\omega_2 = k_1 g_{21} + k_2 g_{22} m + (k_1 g_{22} + k_2 g_{21})\theta.$$

Let G_1 be the matrix corresponding to $(k\omega_1, k\omega_2)$. Then

$$G_1 = k_1 \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} + k_2 \begin{pmatrix} g_{12} m & g_{11} \\ g_{22} m & g_{21} \end{pmatrix}$$
$$= k_1 G + k_2 G\Gamma = G(k_1 I + k_2 \Gamma) = GM(k).$$

Likewise in Case II,

$$k\omega_1 = k_1 g_{11} + k_2 g_{12} m' + (k_1 g_{12} + k_2 g_{11} + k_2 g_{12})\theta,$$
$$k\omega_2 = k_1 g_{21} + k_2 g_{22} m' + (k_1 g_{22} + k_2 g_{21} + k_2 g_{22})\theta$$

so that

$$G_1 = k_1 \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} + k_2 \begin{pmatrix} g_{12} m' & g_{11} + g_{12} \\ g_{22} m' & g_{21} + g_{22} \end{pmatrix}$$
$$= k_1 G + k_2 G\Gamma = G(k_1 I + k_2 \Gamma) = GM(k).$$

If G corresponds to (ω_1, ω_2) and G' to (ω_1', ω_2'), it follows that $GM(k)$ corresponds to $(k\omega_1, k\omega_2)$ and $G'M(k')$ to $(k'\omega_1', k'\omega_2')$. If these latter two ideals are equal, there exists by Theorem 2 an integral matrix C, $|C| = \pm 1$, such that $GM(k) = CG'M(k')$, and conversely. Thus we have proved

THEOREM 5. *A necessary and sufficient condition in order that G_1 and G_2 correspond to ideals of the same class is that there shall exist a relation*

$$G_1 P_1 = CG_2 P_2$$

where P_1 and P_2 are matrices corresponding to integral numbers of F under a Poincaré correspondence, and C is an integral matrix of determinant ± 1.

Suppose that G_1 and G_2 correspond to matrices of the same class. Then there are matrices A_1 and A_2 with integral elements such that

$$G_1\Gamma = A_1 G_1, \qquad G_2\Gamma = A_2 G_2,$$

and matrices P_1 and P_2 corresponding to integral numbers of F such that

$$G_1 P_1 = CG_2 P_2 \qquad |C| = \pm 1.$$

Since P_1 and P_2 correspond to integral numbers of F under Poincaré correspondences, they are commutative with Γ. Thus

$$G_1 P_1 \Gamma = G_1 \Gamma P_1 = A_1 G_1 P_1,$$
$$C G_2 P_2 \Gamma = C G_2 \Gamma P_2 = C A_2 G_2 P_2 = C A_2 C^{-1} C G_2 P_2.$$

Since every matrix considered is non-singular, $A_1 = C A_2 C^{-1}$. Hence if G_1 and G_2 correspond to equivalent ideals, A_1 and A_2 are similar matrices.

Conversely, suppose that we have

$$G_1 \Gamma = A_1 G_1, \qquad G_2 \Gamma = A_2 G_2, \qquad A_2 = C^{-1} A_1 C, \qquad |C| = \pm 1.$$

Since neither $|G_1|$ nor $|G_2|$ is zero, we have

$$G_2 \Gamma G_2^{-1} = A_2 = C^{-1} A_1 C = C^{-1} G_1 \Gamma G_1^{-1} C.$$

Multiply on the left by $G_1^{-1} C$ and on the right by G_2, obtaining

$$G_1^{-1} C G_2 \Gamma = \Gamma G_1^{-1} C G_2.$$

Let S be the scalar matrix each of whose diagonal elements is $|G_1|$. Then $S G_1^{-1} C G_2$ has integral elements and, since it is commutative with Γ, is of the form $k_1 I + k_2 \Gamma = M(k_1 + k_2 \theta) = P$. Then

$$S G_1^{-1} C G_2 = P, \qquad C G_2 S = G_1 P,$$

since the scalar matrix S is commutative with every matrix. Moreover, it corresponds to $|G_1|$ under a Poincaré correspondence, so that the ideals determined by G_1 and G_2 are in fact equivalent. Thus we have

THEOREM 6. *A necessary and sufficient condition in order that the ideals corresponding to G_1 and G_2 be equivalent (belong to the same class) is that the matrices $G_1 \Gamma G_1^{-1}$ and $G_2 \Gamma G_2^{-1}$ be similar. The class number of F is the number of dissimilar matrices which are transforms of Γ.*

We shall henceforth call matrices of the type $A = G \Gamma G^{-1}$ *class matrices.*

7. **Basis for an ideal class.** We have seen that we obtain all the ideals of a class if we determine all the matrices G which transform Γ into a given matrix A. Given

$$G = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

let us determine all matrices G which transform Γ into A, i. e., such that $G \Gamma = A G$. In Case I this yields the equations

$$(5)\qquad \begin{aligned} a &= b\alpha + d\beta, & bm &= a\alpha + c\beta, \\ c &= b\gamma + d\delta, & dm &= a\gamma + c\delta. \end{aligned}$$

Since $|G| \neq 0$, we have also the conditions that the determinants and traces of A and Γ be equal, i. e., $\alpha\delta - \beta\gamma = -m$, $\alpha + \delta = 0$. If we solve the equations on the right for a and c, we find

$$a = d\beta - b\delta, \qquad c = b\gamma - d\alpha$$

which, in view of $\delta = -\alpha$, are precisely the values already found for a and c. Hence if we choose $\alpha, \beta, \gamma, \delta$ any integers satisfying $\delta = -\alpha$, $\alpha^2 + \beta\gamma = m$, b and d arbitrary integers and a and c as given by (5), then we shall obtain a G and a A such that $G\Gamma = AG$.

If we use the values of a and c as found in (5) with δ replaced by $-\alpha$, we see that the most general G is of the form

$$G = \begin{pmatrix} b\alpha + d\beta & b \\ b\gamma - d\alpha & d \end{pmatrix} = dB_1 + bB_2$$

where

$$B_1 = \begin{pmatrix} \beta & 0 \\ -\alpha & 1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} \alpha & 1 \\ \gamma & 0 \end{pmatrix}, \qquad \alpha^2 + \beta\gamma = m.$$

Note that both B_1 and B_2 correspond to canonical bases of ideals.

In Case II we have similarly

$$(5')\qquad \begin{aligned} a + b &= b\alpha + d\beta, & bm' &= a\alpha + c\beta, \\ c + d &= b\gamma + d\delta, & dm' &= a\gamma + c\delta, \end{aligned}$$

where $\alpha + \delta = 1$ and $\alpha\delta - \beta\gamma = -m'$. Solving the right hand equations for a and c, we have

$$a = d\beta - b\delta, \qquad c = b\gamma - d\alpha$$

which are the left hand equations in view of $\alpha + \delta = 1$. The most general G is of the form

$$G = \begin{pmatrix} d\beta + b\alpha - b & b \\ b\gamma - d\alpha & d \end{pmatrix} = dB_1 + bB_2$$

where

$$B_1 = \begin{pmatrix} \beta & 0 \\ -\alpha & 1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} \alpha - 1 & 1 \\ \gamma & 0 \end{pmatrix}.$$

Here again both B_1 und B_2 correspond to canonical bases of ideals.

We may at this point prove the known result which we used in § 6, namely that the norm of every ideal is different from zero. This norm N is the determinant of xB_1+yB_2. In Case I,

$$N = \beta x^2+2\alpha xy-\gamma y^2 = \beta\left(x+\frac{\alpha+\sqrt{m}}{\beta}y\right)\left(x+\frac{\alpha-\sqrt{m}}{\beta}y\right).$$

Since $\alpha^2+\beta\gamma = m$ is not a perfect square, we know that $\beta \neq 0$. In order to have $N = 0$, we should require

$$\beta x+(\alpha+\sqrt{m})\,y = 0 \quad \text{or} \quad \beta x+(\alpha-\sqrt{m})\,y = 0.$$

Since 1 and $\sqrt{m}$ are linearly independent over the rational field, each of these equations implies $x = y = 0$. Since every ideal contains numbers other than zero, this case is excluded.

Likewise in Case II,

$$\begin{aligned} N &= \beta x^2+(2\alpha-1)xy-\gamma y^2 \\ &= \beta\left(x+\frac{(2\alpha-1)+\sqrt{m}}{2\beta}y\right)\left(x+\frac{(2\alpha-1)-\sqrt{m}}{2\beta}y\right). \end{aligned}$$

Just as before, $N = 0$ implies $x = y = 0$.

We now have

THEOREM 7. *The set of all matrices corresponding to the ideals of a given class has a basis composed of two matrices B_1 and B_2 each of which corresponds to the canonical basis of an ideal. That is, the totality of such matrices is given by*

$$xB_1+xB_2$$

where x and y are independent rational integral variables not both zero.

8. **Automorphs.** When we choose a definite matrix A to represent an ideal class, we fix the basic ideals B_1 and B_2. Every matrix xB_1+yB_2, $(x, y) \neq (0, 0)$, transforms Γ into A and hence corresponds to an ideal of the class. But it is presumable that more than one matrix of this type corresponds to the same ideal. For instance, suppose that G and $G' = CG$, $|C| = \pm 1$, both transform Γ into A. We should have

$$G\Gamma = AG, \quad CG\Gamma = ACG, \quad G\Gamma = C^{-1}ACG$$

so that $C^{-1}AC = A$.

Conversely, if $C^{-1}AC = A$, then the two matrices G and CG corresponding to the same ideal are both found in the set xB_1+yB_2.

We shall call a matrix C of determinant ± 1 which is commutative with A an *automorph* of A. We proceed to find all automorphs of A.

LEMMA 1. *A necessary and sufficient condition in order that the rational matrices*

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \qquad M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

A not scalar, be commutative is that $M = uI + vA$ where I is the identity matrix and u and v are rational numbers.

The sufficiency of the condition is evident. To prove the necessary part, we note that $MA = AM$ implies

$$\begin{aligned} a\alpha + b\gamma &= a\alpha + c\beta, & a\beta + b\delta &= b\alpha + d\beta, \\ c\alpha + d\gamma &= a\gamma + c\delta, & c\beta + d\delta &= b\gamma + d\delta. \end{aligned}$$

The first and last equations give merely $b\gamma = c\beta$. The second and third give, respectively,

$$(a-d)\beta = (\alpha-\delta)b, \qquad (a-d)\gamma = (\alpha-\delta)c.$$

These three relations are all given by the proportionality

$$b:\beta = c:\gamma = a-d:\alpha-\delta \equiv v.$$

Evidently v is finite unless $\beta = \gamma = \alpha - \delta = 0$, i. e., unless A is a scalar matrix, in which case of course M is arbitrary. Hence we have

$$b = v\beta, \qquad c = v\gamma, \qquad a - d = v(\alpha - \delta).$$

Let us define $a - v\alpha \equiv u$. Then $a = u + v\alpha$, $d = u + v\delta$ so that

$$M = \begin{pmatrix} u+v\alpha & v\beta \\ v\gamma & u+v\delta \end{pmatrix} = uI + vA. \tag{6}$$

LEMMA 2. *If β, γ and $\alpha - \delta$ are relatively prime integers, the only matrices M of Lemma 1 all of whose elements are integers are given by* (6) *with u and v rational integers.*

Suppose that the elements of M are integers, $b = v\beta$, $c = v\gamma$ and $a - d = v(\alpha - \delta)$ and that β, γ and $\alpha - \delta$ are relatively prime integers, while v is of the form p/q where p and q are relatively prime integers. Then q must be a divisor of β, γ and $\alpha - \delta$, and hence is 1, so that v is an integer. Then $u = a - v\alpha$ is also an integer.

Let us now apply these results to determine the automorphs of our class matrices. In Case I we have A with $\alpha = \delta$, $\alpha^2 + \beta\gamma = m$. If β

and γ have an even common factor, then $m \equiv \alpha^2 \pmod 4$, i. e., m is a quadratic residue modulo 4 and hence is congruent to 0 or 1 modulo 4, which is contrary to assumption. Then if h is a common divisor of β, γ and 2α, it is odd and hence a common divisor of β, γ and α. Hence h^2 is a divisor of m. But we assumed that m had no square factor greater than 1. Hence β, γ and $\alpha - \delta$ are relatively prime, and by Lemma 2 the automorphs of A are given by (5) where u and v are integers.

In Case II we have $\alpha + \delta = 1$, $\alpha^2 + \beta\gamma - \alpha = m' = \frac{1}{4}(m-1)$, $m \equiv 1 \pmod 4$ so that

$$\alpha - \delta = 2\alpha - 1, \qquad m = 4\beta\gamma + (2\alpha - 1)^2.$$

If $h > 1$ were a common factor of β, γ and $2\alpha - 1$, h^2 would be a factor of m, contrary to assumption. Thus this case also comes under Lemma 2.

In order then that $C = uI + vA$ be an automorph of A, it is necessary and sufficient that u and v be rational integers such that $|C| = \pm 1$. This latter condition gives in Case I,

$$\pm 1 = u^2 - v^2\alpha^2 - v^2\beta\gamma = u^2 - v^2 m,$$

and in Case II,

$$\pm 1 = u^2 - v^2\alpha^2 + v^2\alpha + vu - v^2\beta\gamma = u^2 + uv - m'v^2.$$

We have now proved

THEOREM 8. *The totality of automorphs of A is given by $uI + vA$ where u and v are rational integers satisfying Pell's equation*

$$u^2 - v^2 m = \pm 1, \qquad u^2 + uv - v^2 m' = \pm 1.$$

It is well known that the units of F are given by $u + v\theta$ where u and v are the solutions of Pell's equation. When $A = \Gamma$, Theorem 8 merely tells us that two principal ideals $\{d_1\}$ and $\{d_2\}$ are equal if and only if $d_1 = \varepsilon d_2$ where ε is a unit, i. e., that there is a p-to-1 correspondence between the integral numbers of F and the principal ideals, where p is the number (finite or infinite) of units.

Let $A = G\Gamma G^{-1}$ be a class matrix not similar to Γ, and $C_1 = I, C_2, C_3, \cdots$ the automorphs of A. The matrices $C_1 G, C_2 G, C_3 G, \cdots$ all correspond to the same ideal. Thus if B_1 and B_2 are the basic matrices determined by A, the matrices $xB_1 + yB_2$, $(x, y) \neq (0, 0)$, can be arranged into sets of p matrices each, each set corresponding to a unique ideal. We now have

THEOREM 9. *The matrices $xB_1 + yB_2$, $(x, y) \neq (0, 0)$ of Theorem 7 are in p-to-1 correspondence with the ideals of the class whose class matrix is $A = B_1 \Gamma B_1^{-1}$, p being the number of units in the field.*

9. **Reduced class matrices.** To each ideal class there corresponds not merely one class matrix, but an entire set of similar matrices. It would be desirable to be able to define reduced matrices which would be in one-to-one correspondence with the ideal classes. This we shall be able to do for the imaginary fields, but for real fields we shall have to be content with a correspondence between ideals and sets of reduced matrices, each set containing a finite number of matrices.

In order to save a long and intricate calculation of reduced matrices, we shall establish a *liaison* with the classical theory of reduced ideals and merely translate the known results into the new notation.

In every ideal class of an imaginary field there is one and only one *reduced** ideal $[a, (b+\sqrt{-\Delta})/2]$ such that $-\Delta = 4m$ when $m \equiv 2, 3 \pmod 4$ and $-\Delta = m$ when $m \equiv 1 \pmod 4$ such that $b^2+\Delta$ is divisible by $4a$, and either $-a < b \leqq a$, $b^2+\Delta > 4a^2$ or $0 \leqq b \leqq a$, $b^2+\Delta = 4a^2$. We know that b is even when $m \equiv 2, 3 \pmod 4$ and b is odd when $m \equiv 1 \pmod 4$.

In Case I the reduced ideal corresponds to the matrix

$$B = \begin{pmatrix} a & 0 \\ b/2 & 1 \end{pmatrix}$$

where $b/2$ is an integer. The corresponding class matrix is

$$\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} = A = B\Gamma B^{-1} = \begin{pmatrix} -b/2 & a \\ \gamma & b/2 \end{pmatrix}$$

where γ is the integer defined by $b^2 - 4m = -4a\gamma$. Thus $\alpha = -b/2$, $\beta = a$, and the class matrix A is said to be reduced if $\alpha^2+\beta\gamma = m$ and either

$$-\beta \leqq 2\alpha < \beta, \qquad 0 < \beta < -\gamma,$$

or

$$-\beta \leqq 2\alpha \leqq 0, \qquad 0 < \beta = -\gamma.$$

In Case II, $b/2+\sqrt{-\Delta}/2 = (b-1)/2+\theta$ so that the reduced ideal corresponds to

$$B = \begin{pmatrix} a & 0 \\ \dfrac{b-1}{2} & 1 \end{pmatrix}.$$

In this case

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & 1-\alpha \end{pmatrix} = B\Gamma B^{-1} = \begin{pmatrix} \dfrac{1-b}{2} & a \\ \gamma & \dfrac{1+b}{2} \end{pmatrix}$$

* H. H. Mitchell, *Ann. of Math.*, vol. 27 (1926), p. 300.

where γ is defined by $b^2 - m = -4a\gamma$. Thus $\alpha = (1-b)/2$, $\beta = a$, and $\mathcal{A}$ is said to be reduced if $(2\alpha-1)^2 + 4\beta\gamma = m$ and either

$$-\beta \leqq 2\alpha - 1 < \beta, \qquad 0 < \beta < -\gamma$$

or

$$-\beta \leqq 2\alpha - 1 \leqq 0, \qquad 0 < \beta = -\gamma.$$

We have so defined reduced class matrix Δ that when $\mathcal{A}$ is reduced the ideal matrix B_1 of Theorem 7 corresponds to a reduced ideal in the classical sense. Note that $\mathcal{A}$ and B_1 determine each other uniquely. Since every class contains one and only one reduced ideal, it follows that every class matrix is similar to one and only one reduced class matrix. Hence we have

THEOREM 10. *In an imaginary quadratic field every class matrix is similar to one and only one reduced class matrix, so that the number of reduced class matrices is the class number of the field. When the class matrix is reduced the corresponding ideal matrix B_1 of Theorem 7 corresponds to a reduced ideal in the classical sense.*

10. **Real fields.** When $m > 0$, a reduced ideal according to the usual definition* is one which possesses a canonical basis $[a, (b+\sqrt{D})/2]$ where $b^2 - D \equiv 0 \pmod{4a}$ and $0 < \sqrt{D} - b < 2a < \sqrt{D} + b$. Here $D = 4m$ and b is even in Case I, and $D = m$ and b is odd in Case II.

In Case I the matrix corresponding to a reduced ideal is

$$B' = \begin{pmatrix} a & 0 \\ b/2 & 1 \end{pmatrix},$$

and

$$\mathcal{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} = B'\Gamma B'^{-1} = \begin{pmatrix} -b/2 & a \\ \gamma & b/2 \end{pmatrix}$$

where $b^2 - 4m = -4a\gamma$. We shall then say that $\mathcal{A}$ is reduced if $\alpha^2 + \beta\gamma = m$ and $0 < \sqrt{m} + \alpha < \beta < \sqrt{m} - \alpha$.

In Case II, the matrix corresponding to a reduced ideal is

$$B' = \begin{pmatrix} a & 0 \\ \dfrac{b-1}{2} & 1 \end{pmatrix}$$

so that

$$\mathcal{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & 1-\alpha \end{pmatrix} = B'\Gamma B'^{-1} = \begin{pmatrix} \dfrac{1-b}{2} & a \\ \gamma & \dfrac{1+b}{2} \end{pmatrix}$$

* H. H. Mitchell, loc. cit., p. 302.

where γ is defined by $b^2 - m = -4a\gamma$. We shall say that A is reduced if $(2\alpha-1)^2+4\beta\gamma = m$ and $0<\sqrt{m}+(2\alpha-1)<2\beta<\sqrt{m}-(2\alpha-1)$. Again in each case B' is the matrix B_1 of Theorem 7 when A is reduced.

While it is true that every class contains at least one reduced ideal and therefore every class matrix is similar to a reduced class matrix, it is not true that a class ordinarily contains but one reduced ideal. There is a definite procedure* however, by which one may separate the reduced ideals into cycles, each cycle corresponding to an ideal class. Translated into our present notation, the process is as follows:

Apply to a reduced class matrix A the transformation CAC^{-1} where

$$C = \begin{pmatrix} l & 1 \\ 1 & 0 \end{pmatrix}$$

and l is defined in the respective cases as the integer satisfying

$$l = \frac{\alpha-\sqrt{m}}{\beta}+\varepsilon, \qquad l = \frac{(2\alpha-1)-\sqrt{m}}{2\beta}+\varepsilon, \qquad 0<\varepsilon<1.$$

Just as in the usual treatment of ideals, it follows that $A' = CAC^{-1}$ is also reduced, and that all reduced ideals of a given class are obtained by starting with any one reduced matrix and applying the above transformation successively until the original matrix appears again. Hence we have

THEOREM 11. *In a real quadratic field every class matrix is similar to one and only one cycle of reduced class matrices, so that the number of cycles is the class number of the field. When a reduced class matrix is used, the corresponding ideal matrix B_1 of Theorem 7 corresponds to a reduced ideal in the classical sense.*

11. **Conclusions.** When we pass from quadratic integers to principal ideals, our numbers undergo a condensation, for if $u_1, u_2, \cdots, u_p$ (where p is finite or infinite) denote the units of the field, then all the integral numbers $u_1x, u_2x, \cdots, u_px$, yield the same principal ideal $\{x\}$. In like manner when we pass from the ideal matrices xB_1+yB_2 of any class to the ideals to which they correspond, we experience a similar condensation, for if X is one such matrix, all the matrices $C_1X, C_2X, \cdots, C_pX$, where C is an automorph of $X\Gamma X^{-1}$, correspond to the same ideal. It might then seem desirable to think of the existence of certain *ideal numbers* (reserving the term "ideal" for the older concept) which bear the same relation to non-principal ideals as the quadratic integers do to principal

* H. H. Mitchell, loc. cit., p. 303.

ideals. We may think of these ideal numbers as numbers corresponding under a Poincaré correspondence to non-principal ideal matrices, or indeed we may think of them as the matrices themselves. To each class of ideal numbers we may adjoin zero and the units $uI+vA$ where u and v satisfy Pell's equation. Thus each ideal class has the same numbers p of units. Within the same class ideal numbers (not units) may be added uniquely, since they are all given without repetition by the formula xB_1+yB_2 where x and y are rational integer.

If two matrices correspond to principal ideals, their matric product corresponds to the product of the ideals. It would be a very interesting result if this were true for all ideals, but unfortunately it is not. Ideal multiplication does not coincide with matric multiplication in non-principal classes, and addition of ideal matrices is not usually distributive with respect to ideal multiplication. Poincaré may have had this fact in mind, but it is not evident from a mere reading of his notes, which leave the reader with the hope that the theory of ideals may be reduced to the study of matric algebras.

OHIO STATE UNIVERSITY.

A CLASS OF FUNCTIONAL EQUATIONS.*

By Einar Hille.

1. The author has previously† studied the functional equation

$$|F(x+iy)|^2 = |F(x)|^2+|F(iy)|^2. \tag{1}$$

This equation is evidently a special case of the following general problem: Given an analytic function $R(u, v)$ of two variables u and v; is it possible to determine an analytic function $f(z)$ such that

$$|f(x+iy)| = R\{|f(x)|, |f(iy)|\}\,? \tag{2}$$

We shall assume that $R(u, v)$ is a rational function of u and v. The extension to the case in which $R(u, v)$ is an algebraic function does not introduce any additional difficulties, but will not be considered here. In §§ 4 and 5 of the present paper we indicate a general method of attacking this problem which will give the solution whenever a solution exists. The functions $f(z)$ which are defined in this manner are as a rule infinitely many-valued; assuming $R(u, v)$ to be symmetric, we determine in § 6 the possible single-valued solutions. The latter solutions are either of the form $\exp(c_1 z^2+c_2 z+c_3 i)$ or are elliptic functions of simple type including degenerate cases of such functions. The analysis is carried through in a couple of illustrating cases in § 7.

2. It is obvious that the function $R(u, v)$ cannot be a perfectly arbitrary rational function. To begin with, $R(u, v)$ must have real coefficients and a region should exist in the real uv-plane in which $R(u, v)>0$. Before we go any further let us dispose of the trivial case in which $R(u, v)$ depends upon only one variable, u for instance. Then we have

$$w = R(u). \tag{3}$$

If we let $y\to 0$, $w\to u$ and $u = R(u)$. This equation may be satisfied identically or only for discrete values of u. In the former case the functional equation becomes

$$w = u, \tag{4.1}$$

* Received October 11, 1927; presented to the American Mathematical Society, October 28, 1922.

† A Pythagorean Functional Equation, Annals of Math., (2), vol. 24, no. 2, pp. 175–180, 1922.

the solution of which is given by

$$f(z) = e^{c_1 z + c_2} \tag{4.2}$$

where c_1 and c_2 are constants, the former real the latter imaginary.

In the latter case the equation $u = R(u)$ must have a certain number of real positive roots $u_1, u_2, \cdots, u_n$, if the problem shall have a solution. The corresponding solutions are

$$f_\nu(z) = u_\nu e^{i\theta_\nu} \qquad (\nu = 1, 2, \cdots, n)$$

where θ_ν are arbitrary real constants.

3. From this point on we shall assume that $R(u, v)$ depends effectively upon two variables u and v. In order to simplify the discussion we shall make the restrictive assumption that $R(u, v)$ is a symmetric function of u and v.

Let us further assume that the solution $f(z)$ approaches definite limiting values when z tends towards the origin along either of the axes and let us put

$$\lim_{y\to 0} |f(iy)| = a, \qquad \lim_{x\to 0} |f(x)| = b \tag{5}$$

where $0 \leqq a, b \leqq +\infty$. Then

$$u = R(u, a) \quad \text{and} \quad v = R(b, v). \tag{6}$$

If a and b are finite numbers they satisfy the algebraic equation

$$t = R(t, t). \tag{7}$$

By considering all the different roots of this equation which are real and $\geqq 0$, we obtain all possible values for a and b. We may lose an infinite root but if such a root exists the equation

$$\frac{1}{s} = R\left(\frac{1}{s}, \frac{1}{s}\right)$$

has the root zero. Let a be a root of (7). There are three different possibilities to consider, namely

$$\text{(8)} \quad \text{(i) } R(u, a) \equiv u, \qquad \text{(ii) } R(u, a) \equiv a, \qquad \text{(iii) } R(u, a) \not\equiv u, a.$$

The first case is the interesting one for our purpose. We shall show that there exists only one such critical value. Suppose there were two values, a and b. Then

$$u = R(u, a), \qquad u = R(u, b). \tag{9}$$

Hence

$$a = \lim_{u \to a} R(u, b) = R(a, b),$$
$$b = \lim_{u \to b} R(u, a) = R(b, a).$$

In view of the symmetry of $R(u, v)$ in u and v these expressions are equal, consequently $a = b$.

We can always assume that the critical value a is either 0 or 1. In fact, if a has a finite value, c, different from 0 or 1 we can replace the given functional equation (2) by

$$(2^*) \qquad W = \frac{1}{c} R(cU, cV) = R_1(U, V)$$

which admits of $a = 1$ as critical value. If $f(z)$ is the general solution of the new functional equation, $c f(z)$ is the general solution of the old one. On the other hand, if the critical value is infinite we replace u, v, w by $\frac{1}{u}, \frac{1}{v}, \frac{1}{w}$ respectively. The resulting equation has 0 as critical value.

We notice that if $v = a$ then

$$(10) \qquad R(u, v) \equiv u, \quad \frac{\partial}{\partial u}[R(u, v)] = 1, \quad \frac{\partial^2}{\partial u^2}[R(u, v)] = 0.$$

The second set of values defined by (8) we call singular values. They correspond to singular constant solutions. In fact, suppose that

$$(11.1) \qquad R(u, s) \equiv s$$

where $s \geqq 0$ is some constant. Then the functional equation has the solution

$$(11.2) \qquad f(z) = s e^{i\theta}.$$

The third set of values has no particular connection with our problem.

4. We shall outline a general method of attacking the problem. We assume that $R(u, v)$ is positive in a certain region of the uv-plane, e. g. when $u \geqq c$, $v \geqq c$, and that a value $a \geqq c$ exists such that $R(u, a) \equiv u$. The hypothesis that $R(u, v)$ is symmetric in u and v is not necessary for the discussion immediately following but will be used later.

Suppose that $f(x + iy)$ is an analytic solution of (2), and consider any simply-connected region within which $f(z)$ is holomorphic and different from zero. In such a region $\log |f(x + iy)|$ is a harmonic function and satisfies Laplace's equation:

$$(12.1) \qquad \Delta \log |f(x + iy)| = 0$$

or

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\log R(u,v)=0. \tag{12.2}$$

The equation (12.2) is found to be a mixed differential equation in two independent variables x and y and two dependent variables u and v, namely

$$R\frac{\partial R}{\partial u}\frac{d^2u}{dx^2}+R\frac{\partial R}{\partial v}\frac{d^2v}{dy^2}+\left[R\frac{\partial^2 R}{\partial u^2}-\left(\frac{\partial R}{\partial u}\right)^2\right]\left(\frac{du}{dx}\right)^2 + \left[R\frac{\partial^2 R}{\partial v^2}-\left(\frac{\partial R}{\partial v}\right)^2\right]\left(\frac{dv}{dy}\right)^2=0. \tag{13}$$

If $u=g(x)$ and $v=h(y)$ form a solution of this differential equation, then $u=g(\alpha x+\beta)$ and $v=h(\alpha y+\gamma)$ also form a solution and that for arbitrary values of α, β and γ. Not all these solutions of the differential equation can be solutions of the functional equation. In fact, assuming that $R(u,v)$ is symmetric in u and v, we have $u=a$ and $v=a$ for x and y respectively equal to zero. This implies two conditions on the constants. Further, we have the requirement that $u\geqq 0$ and $v\geqq 0$ for real values of x and y respectively. This imposes still another restriction upon the constants.

It may or may not be possible to fulfill these conditions. Suppose for the moment that we can find a pair of functions

$$u=G(x), \quad v=H(y),$$

satisfying equation (13) and such that

(1) $G(x)\geqq 0$, $H(y)\geqq 0$ when x and y are real;

(2) $G(0)=a$, $H(0)=a$;

(3) $R(G(x), H(y))>0$ when $-x_1\leqq x\leqq +x_1$, $-y_1\leqq y\leqq +y_1$.

Then

$$P(x,y)=\log R(G(x), H(y))$$

is a harmonic function in the rectangle indicated. The conjugate harmonic function $Q(x,y)$ is determined uniquely save for an arbitrary additive constant. Then form

$$f(x+iy)=e^{P(x,y)+iQ(x,y)}.$$

This function is a solution of the functional equation.

5. If we keep y (or x) constant in equation (13) we obtain an ordinary differential equation which can be integrated by means of 3 quadratures.

The most favorable choice which we can make for a constant value of y is 0. Then $v = |f(iy)| = a$. Suppose that

$$\frac{dv}{dy} \to s, \quad \frac{d^2v}{dy^2} \to t \text{ when } y \to 0. \tag{14}$$

If a equals zero we have $s = 0$, but it may happen that $s = 0$ for other values of a as well. Then (13) becomes

$$u\frac{d^2u}{dx^2} - \left(\frac{du}{dx}\right)^2 + P(u) = 0 \tag{15}$$

where

$$P(u) = tu\frac{\partial R}{\partial v}\Big]_{v=a} + s^2\left[u\frac{\partial^2 R}{\partial v^2} - \left(\frac{\partial R}{\partial v}\right)^2\right]_{v=a}. \tag{16}$$

Let us introduce a new variable

$$U = \left(\frac{du}{dx}\right)^2. \tag{17.1}$$

Then

$$\frac{1}{2}u\frac{dU}{du} - U + P(u) = 0 \tag{17.2}$$

or

$$U = u^2\left[C - \int_a^u \frac{P(u)}{u^3}\,du\right]. \tag{17.3}$$

This expression has to be modified if $a = 0$ owing to the singularity of the integrand at the origin. Hence

$$x = \int_a^u v^{-1}\left[C - \int_a^v \frac{P(\tau)}{\tau^3}\,d\tau\right]^{-\frac{1}{2}} dv. \tag{18}$$

Since $u = |f(x)|$ and $v = |f(iy)|$ we conclude that if $f(z)$ is holomorphic at the origin

$$u(0) = v(0) = a, \qquad \left[\frac{du}{dx}\right]_{x=0} = \left[\frac{dv}{dy}\right]_{y=0} = s. \tag{19}$$

Hence if $a \neq 0$ we have $C = s^2 a^{-2}$. If $a = 0$ we have to make appropriate changes in the lower limit of integration in (17.3) and (18).

The expression for y is obtained by changing u into v in the formulas above. The expressions for u and v are obtained by inversion from formula (18) and the corresponding formula for y. If this process leads to a pair of functions having the properties specified at the end of § 4,

we have obtained a particular solution of the problem from which the general solution is easily obtained.

6. We are now in position to consider the problem of determining when the functions $u(x)$ and $v(y)$ and consequently also $f(z)$ will be single-valued functions of their arguments. Painlevé* has investigated the differential equations of the second order

$$y'' = F(y', y, x)$$

which have fixed critical points. The right hand side of the equation is supposed to be rational in y', algebraic in y and analytic in x. In particular, the equation

$$y'' = \frac{y'^2}{y} + a(y), \tag{20.1}$$

where $a(y)$ is an algebraic function of y not depending upon x, has a single-valued solution only if $a(y)$ has either of the following two forms, namely

$$a(y) = \beta y, \tag{20.2}$$

or

$$a(y) = \alpha y^3 + \beta y^2 + \gamma + \frac{\delta}{y}. \tag{20.3}$$

In the first case the solutions have the form

$$y = Ke^Y, \qquad Y = \frac{\beta}{2}x^2 + K_1 x. \tag{21}$$

In the second case the solutions are elliptic functions of simple type or degenerate cases of such functions.

Hence, if a functional equation of the type (2) has a single-valued solution, this solution must be either an exponential function of the type represented by formula (21), or an elliptic function or a degenerate case of such a function. The elliptic functions which satisfy equation (20.1) with (20.3) and in addition are real upon the axes form a very restricted class. On account of the conditions of reality the lattice points $2n\omega_1 + 2m\omega_3$ where $2\omega_1$ and $2\omega_3$ are the primitive periods, must be symmetrically distributed with respect to the axes. Hence the primitive periods are either conjugate imaginary or one is real and the other purely imaginary.

* Acta Math., vol. 25, pp. 1–85, 1902. The results which we need are to be found in Tableau II, p. 24, compared with Tableau XI, equations (13) and (15), p. 55. The discussion of Painlevé was completed in one important point by Gambier, Acta Math., vol. 33, 1910, pp. 1–55. The latter's results are not needed for our purposes.

7. We shall study a couple of simple examples. We begin with the equation

$$w = uv. \tag{22}$$

There are two critical values namely 0 and 1. The former is a singular value, the latter not. The singular solution is identically zero. In order to find the other solution we form the differential equation (13) corresponding to the problem. This equation can be written in the form

$$\frac{d^2}{dx^2}[\log u] + \frac{d^2}{dy^2}[\log v] = 0. \tag{23}$$

The first term is a function of x alone, the second one of y alone. Hence each term separately must be equal to a constant. Integrating the equation

$$\frac{d^2}{dx^2}[\log u] = C, \tag{24}$$

we find that

$$u = e^{\frac{1}{2}Cx^2+C_1x+C_2},$$
$$v = e^{-\frac{1}{2}Cy^2+C_3y+C_4}.$$

Since $u(0) = v(0) = 1$ we have $C_2 = C_4 = 0$. Further, the remaining constants must be real. Hence

$$f(z) = e^{\frac{1}{2}Cx^2+Kz+K_1 i} \tag{25}$$

where C and K_1 are real constants and K a complex one, otherwise arbitrary. This case corresponds to $a(y) = Cy$ the solution of which was given in formula (21).

Now let us take the equation

$$w = \frac{u+v}{1+uv}. \tag{26}$$

Here again we have the critical values 0 and 1 of which the latter is singular. To this latter value corresponds an arbitrary unit vector as a singular solution. The value $a = 0$ corresponds to an elliptic function. In this case the differential equation (13) reads

$$\begin{aligned}&(u+v)(1+uv)(1-v^2)u'' - [(1+uv)^2+v^2(u+v)^2](u')^2\\ &+(u+v)(1+uv)(1-u^2)v'' - [(1+uv)^2+u^2(u+v)^2](v')^2 = 0.\end{aligned} \tag{27}$$

If we put $y = 0$ we have $v_0 = 0$, $v_0' = 0$. Let us put $v''(0) = 2b^2$. Then the equation takes on the form

$$uu'' - (u')^2 + 2b^2 u(1 - u^2) = 0. \tag{28}$$

In order to integrate this equation we put $u = U^2$ and obtain

$$UU'' - (U')^2 + b^2(1 - U^4) = 0.$$

Multiplying by $2\frac{U'}{U^3}$ and integrating we obtain

$$(U')^2 = b^2(U^4 - cU^2 + 1). \tag{29}$$

Let us determine k from the relation

$$k + \frac{1}{k} = c \tag{30}$$

and put $U = \sqrt{k}\, t$. Then we find that

$$t = \operatorname{sn} \operatorname{am}\left(\frac{bx}{\sqrt{k}} \middle/ k\right)$$

or

$$u = k \operatorname{sn}^2 \operatorname{am}\left(\frac{bx}{\sqrt{k}} \middle/ k\right). \tag{31}$$

Finally

$$f(z) = e^{i\theta} k \operatorname{sn}^2 \operatorname{am}\left(\frac{bz}{\sqrt{k}} \middle/ k\right). \tag{32}$$

Since c is a real constant we find that k is either real or lies on the unit-circle; b is either real or purely imaginary; θ is a real arc.

The special case $k = \pm 1$ gives

$$u = \tan^2(bx) \tag{33}$$

and

$$f(z) = e^{i\theta} \tan^2(bz) \tag{34}$$

where b is either real or purely imaginary.

DEFINITIONS OF ABSTRACT GROUPS.*

By G. A. Miller.

Since the term abstract group is restricted to the laws of combination of operators it is clear that no such group could have been completely defined before the laws of combination relating to mathematical operations were explicitly recognized. It is an interesting fact that such fundamental laws as the commutative, associate, and distributive laws attracted so little attention before the nineteenth century that they failed to receive special names. In particular, the technical terms commutative and distributive seem to have been introduced into the mathematical literature by F. J. Servois, *Annales de Mathématiques,* volume 5 (1814), page 93. It should, however, be emphasized that a proof of the commutative law as regards certain multiplications appears already in the *Elements of Euclid,* Book 7, proposition 16, and that the first proposition of Book 2 of the same work is devoted to the distributive law as regards multiplication and addition. The associate law in multiplication was explicitly recognized in the well known work by A. M. Legendre, entitled *Essai sur la théorie des nombres,* 1798, page 10, but it did not then have a special name.

It is questionable whether the mathematics of the nineteenth century exhibits any more profound step in advance than the one involved in the increased emphasis on the fundamental laws governing the combination of mathematical elements. While the commutative law and the distributive law were thus named by F. J. Servois as early as 1814, the associative law seems to have received a special name at a considerably later date. Such a name was given to it by W. R. Hamilton in a paper read November 13, 1843, and published in the *Proceedings of the Royal Irish Academy,* volume 2 (1844), page 430. This law is especially important in the theory of abstract groups, and it is interesting to note that the first definition of an abstract non-abelian group appeared at the same time as the special name for this law. A rapidly growing emphasis on the laws involved in the combination of mathematical elements is noticable in the literature of the first half of the nineteenth century. Since the distributive law relates to two distinct operations it is not directly involved in the ordinary definitions of an abstract group, which does not concern itself with the nature of the operations which may by involved in its elements themselves, but only with the laws governing the combination of these

* Received October 17, 1927.

elements regarded as single operations. The fact that technical terms for these fundamental laws were introduced so long after these laws were explicitly used throws light on the relation between the use of technical terms and the concepts involved therein.

The most elementary abstract group is cyclic and it is now known that there is one and only one such group of every order. When this order is finite this group may be defined by the fact that it is generated by a single operator, but when it is infinite such a definition is impossible since every infinite group has at least two independent generators. There is no upper limit to the possible number of such generators of an infinite group as results directly from the group composed of all the positive rational numbers when they are combined by multiplication. Moreover, when an infinite group has only two independent generators it is not necessarily cyclic since the dihedral group of every possible order has two and only two independent generators. An abstract definition of every possible cyclic group is that it is composed of the positive and negative powers of a single operator. Since the exponents are combined by addition when the operators are combined by multiplication we have in the cyclic group the fundamental principle of logarithms, which was partially recognized as regards positive numbers by Archimedes, and is found already in a less explicit form in the *Elements* of Euclid.

The ancient Greeks came very close to a definition of an abstract cyclic group when they said that unity is the source of all numbers. The idea involved in this statement was doubtless that unity generates all the natural numbers with regard to the operation of addition. If the ancient Greeks had extended their number concept so as to include zero and negative numbers they would have seen in their integers the infinite cyclic group whose independent generators are 1 and -1 when these integers are combined by addition. We have here an early illustration of the simplification effected by the concept of group since the operations of addition and subtraction coalesce when we consider the totality of integers as a group with respect to the operation of addition, but the ancient Greeks did not effect this simplification, and hence they did not reach completely a definition of the concept of the abstract cyclic group. The development of this definition extended through a period of more than a thousand years, and when it was finally completed at about the beginning of the nineteenth century the explicit use of the concept congruence of numbers with respect to a given modulus presented another definition of the concept of cyclic group. The introduction of the concept of modulus was especially due to the work of C. F. Gauss, who also gave properties of the generators of the cyclic groups formed by the n nth roots of unity, these generators being the

primitive roots. Their number, now known as the totient of n, had been determined somewhat earlier by L. Euler.

According to what was noted above the abstract cyclic group is represented by a set of integers when they are combined by addition. The latter group is actually concrete since its elements are numbers, but since in the definition of an abstract cyclic group we must consider powers of an element and the indices of these powers may be represented by integers it is clear that with every abstract cyclic group there is commonly associated a group whose elements are numbers. Hence integers are a usual concomitant of the notion of abstract cyclic group. It should, however, be noted that the elements of the group itself are not necessarily these integers. In abstract group theory we do not only use symbols to respresent the elements, and these symbols are obviously concrete but we also use integers to denote powers of these elements. Hence the term abstract group is a relative term.

While a definition of the term abstract cyclic group was developed very slowly during a period of more than a thousand years, definitions of various other abstract groups appeared in their modern form. The oldest definition of an abstract group after the cyclic group seems to be that of the quaternion group, which was defined in 1843 by W. R. Hamilton by means of the well known symbols i, j, k. Any pair of these symbols generates this group as results directly from their laws of combination, which were given by W. R. Hamilton. About thirteen years later the same author presented to the Royal Irish Academy a paper entitled *A new system of roots of unity*, in which he considered the equations

$$1 = i^2 = k^3 = l^r, \; ik = l, \; r = 3, \; 4, \; 5.$$

He observed that these equations are sufficient to restrict the elements represented by i and k as regards multiplication under the associative law in every possible manner, so that the total number of the distinct products may in each case be interpreted as movements of a regular polyhedron. These beautiful abstract definitions of the groups of the Platonic solids present a great contrast when they are compared with the abstract definition of the quaternion group. While the laws of combination involved in the former illustrated a very old subject and must have been known, at least in part, at a very early date, there is no evidence that any one before W. R. Hamilton knew of any physical object which illustrates the laws of combination involved in the quaternion group.

In the definitions of abstract groups which have been noted thus far no direct reference was made by their authors to the subject of group

theory even if this subject in the form of substitution groups had been developed to a considerable extent before the publication by W. R. Hamilton which were noted above. About two years before the latter of these two publications A. Cayley began to publish articles which are explicitly devoted to the theory of abstract groups. In the first of these articles, which appeared in the *Philosophical Magazine*, volume 7 (1854), page 408, he gave definitions of various fundamental abstract groups, including the non-cyclic group of order 4 and the non-cyclic group of order 6, and he stated correctly that there are only two abstract groups of the latter order. This statement is the more interesting in view of the fact that about 24 years later he stated incorrectly that there are three abstract groups of this order, in his first article published in the *American Journal of Mathematics*, volume 1 (1878), page 51. This fact is important in the history of abstract group theory since it illustrates the slow progress that was made along this line during the first quarter of a century after A. Cayley began to publish on this subject.

While we find in these early writings definitions of abstract groups which are now entirely satisfactory it must be admitted that they were less satisfactory at the time when they were first published since the general term abstract group had not then been defined in a satisfactory manner. Such definitions began to appear after the years 1880, and hence all the earlier definitions of special abstract groups, or of special abstract systems of groups, lack definiteness. On the other hand, A. Cayley came close to giving a definition of the general term abstract group when he assumed in his first article on the subject both the associative law and the fundamental fact that when all the operators of a group are multiplied on the same side by a particular operator thereof the group is reproduced. This implies the postulate that when any two of the three symbols in the equation

$$s_\alpha s_\beta = s_\gamma$$

are contained in a group then the third is also contained therein and is completely determined by this equation. Hence A. Cayley had, as early as 1854, the modern idea of a definition of an abstract group but he failed to state explicitly what assumptions are involved in such a definition, and he does not seem to have been interested in finding the simplest set of such assumptions. In fact, he seems to have overlooked at first the need of a postulate in regard to the uniqueness of a solution.

Since the dihedral groups are the groups of movements of the regular polygons one might have expected that their abstract properties would have been noted at an early date. It is now known that every pair of

operators of order 2 generates a dihedral group whose order is twice the order of the product of these two operators, and that there is one and only one abstract dihedral group of every even order. This definition of the abstract dihedral groups seems not to have been explicitly stated until quite recently*. A. Cayley gave another definition of the abstract dihedral group at an earlier date, viz., that it is generated by two operators s_1, s_2 of orders n and 2 respectively which satisfy the equation

$$s_1 s_2 = s_2 s_1^{n-1}$$

and he noted that there is such a group whose order is any even number.† His observations are equivalent to saying that if an operator of order 2 transforms an operator of order n into its inverse then these two operators generate the dihedral group of order $2n$. He considered in this article also the more general question of groups which are generated by two operators s_1, s_2 when one of these operators transforms the other into a power of itself, and he noted that all the operators of such a group are involved in each of the following two expressions:

$$s_1^p s_2^q, \quad s_2^p s_1^q.$$

The most elementary dihedral group is the non-cyclic group of order 4 which was defined in 1854, as an abstract group, by A. Cayley. This is also the only abelian dihedral group, and it was noted by J. Royce that it represents the one group operation which the Boolean calculus in its original form permits.‡ It is known as the axial group since it is simply isomorphic with the group formed by rotations through π around the three axes of rectangular system of coördinates. It is also known by a number of other names, and it is remarkable that an abstract definition of such a widely useful group was not formulated at an earlier date. The lateness of this formulation constitutes further evidence of the fact that the formal laws of the combination of mathematical operators were practically unnoticed in the mathematical developments preceding those of the last hundred years. It is now difficult to determine the motives which led to a formulation of these laws. Probably the theory of substitution groups exerted a powerful influence in stressing them, but it is not certain that this theory affected the work of F. J. Servois along this line. On the other hand, his work was directly affected by the concept of function,

* G. A. Miller, *Bull. Amer. Math. Soc.*, vol. 7 (1901), p. 424.

† A. Cayley, *Phil. Mag.*, vol. 18 (1859), p. 37.

‡ J. Royce, *Journ. of Phil., Psych.* and *Sci. Methods*, vol. 10 (1913), p. 618.

and hence it appears that the idea of function influenced the development of the concept of an abstract group.

While definitions of abstract groups frequently embody properties of known groups and aim merely to select such properties as are sufficient to define these groups, there are also many instances where definitions were proposed before it was known what groups, or what categories of groups, are defined thereby. This was the case, for instance, as regards the generalizations of the groups of the Platonic solids. Since there is no upper limit to the number of such definitions it is clear that there is here an inexhaustible field for scientific work, where the interest naturally increases with the increase of the number of known fundamental facts relating thereto. Even the groups which can be generated by two operators present many difficulties which remain unsolved.

The determination of all the possible abstract groups of given orders obviously implies a definition of the groups. We have already referred to the fact that in 1854 A. Cayley determined the two possible groups of order 6. In the same article he also noted, what is frequently credited to later writers, that there is only one group of every prime order, and he gave the two possible groups of order 4, but he did not state explicitly that there is no other group of this order. In 1859 he determined the five possible groups of order 8, but the five possible groups order 12 were first determined completely thirty years later. In 1886 A. B. Kempe had stated that there are five groups of this order and gave what he supposed to be a complete determination thereof, but A. Cayley noted later that this determination is incorrect. The dicyclic group does not appear in A. B. Kempe's list, and the substitutions which appear in its place do not constitute a group. The quaternion group is also missing in A. B. Kempe's list of groups of order 8 notwithstanding the fact that A. Cayley had included it in his earlier list of groups of this order. The groups of order p^2 and of order pq, p and q being prime numbers, were first published in Netto's *Substitutionentheorie*, 1882, page 133. After 1890 definitions of various abstract groups began to appear at a much more rapid rate and this date may therefore be selected as marking the close of the early definitions of abstract groups.

CONCERNING A SET OF METRICAL HYPOTHESES FOR GEOMETRY.*

By J. L. Dorroh.

In his paper, Sets of metrical hypotheses for geometry,† R. L. Moore raises the question whether from the set O of order axioms and the set C of congruence axioms employed therein it follows that every segment has a mid-point. It is the purpose of the present paper to show that this question may be answered in the affirmative.

THEOREM 1. *In a given plane containing a line there is not more than one perpendicular‡ to that line at a given point.*

Proof. Suppose that OA, OB and OK are coplanar lines and that OA and OK are perpendicular to the line OB. By Axiom C_1 of M. H. there exists a point C such that BOC§ and $BO \equiv OC$. Since $\sphericalangle AOB \equiv \sphericalangle AOC$, $AO \equiv AO$ and $BO \equiv OC$, therefore, by Theorem 11 of M. H. § 1, $AB \equiv AC$. The line OK is not identical with the line OB, therefore, by the Corollary to Theorem 16 of S. A., the line OK contains a point D such that D is A, ADB or ADC. Suppose ADC. Let E be a point such that DEO. Since the line OK is perpendicular to the line OB, $EB \equiv EC$. We have $AB \equiv AC$, $EB \equiv EC$ and $AE \equiv AE$. It follows from Definition 3 of M. H. § 1 that $\sphericalangle BAE \equiv \sphericalangle CAE$. Since ADC and DEO, therefore by Theorem 7 of S. A. there exists a point F such that AEF and CFO. Since $\sphericalangle BAF \equiv \sphericalangle CAF$, $AB \equiv AC$ and $AF \equiv AF$, therefore $BF \equiv CF$. Thus BC has two distinct mid-points O and F. This contradicts Theorem 23 of M. H. § 1. The assumption ADB leads to the same contradiction. Hence D is A and the line OK is identical with the line OA.

THEOREM 2. *If A, B, C are three non-collinear points and $\sphericalangle ABC \equiv \sphericalangle BAC$ then $AC \equiv BC$.*

Proof. By Theorem 8 of M. H. § 1 either $AC \equiv BC$, $AC > BC$ or $AC < BC$. Suppose $AC > BC$ and let D be a point on the ray BC such

---

* Received September 24, 1927; presented to the American Mathematical Society September 9, 1927.

† Trans. Amer. Math. Soc. vol. 9 (1908), pp. 487–512. This paper will be referred to as M. H. The notation S. A. will be used to denote O. Veblen's paper, A system of axioms for geometry, ibid., vol. 5 (1904), pp. 343–384.

‡ Cf. Definition 7 of M. H. § 3.

§ Used as a sentence, "ABC" means A, B and C are in the order ABC.

that $AC \equiv BD$. Then $BD > BC$ and BCD.* Since $\sphericalangle ABD \equiv \sphericalangle BAC$, $AB \equiv AB$ and $AC \equiv BD$, therefore $AD \equiv BC$. Hence $\sphericalangle ABC \equiv \sphericalangle BAD$. But $\sphericalangle ABC \equiv \sphericalangle BAC$. Therefore, by Theorem 14 of M. H. § 1, $\sphericalangle BAD \equiv \sphericalangle BAC$. Since $AC > AD$, there exists a point E such that ADE and $AC \equiv AE$.

Since ADE and BCD, there exists a point F such that ECF and AFB. Since $\sphericalangle FAC \equiv \sphericalangle FAE$, $AC \equiv AE$ and $AF \equiv AF$, therefore $FC \equiv FE$. Hence E is C and D is C since D is the intersection of the lines BC and AE. Hence $AC \equiv BC$. The assumption that $AC < BC$ leads to the same conclusion.

THEOREM 3. *If A, B, C are three non-collinear points, there exists, within† $\sphericalangle BAC$, a point O such that $\sphericalangle BAO \equiv \sphericalangle OAC$.*

Proof. Let B' denote a point on the ray AC such that $AB \equiv AB'$. Let D and D' denote points such that ADB, $AD'B'$ and $AD \equiv AD'$ and $DB \equiv D'B'$.‡ Since, by Theorem 12 of M. H. § 1, $\sphericalangle BAB' \equiv \sphericalangle BAB'$, therefore $BD' \equiv B'D$ and $\sphericalangle DB'B \equiv \sphericalangle D'BB'$. Since ADB and $AD'B'$, therefore, by Theorem 13 of S. A., there exists a point O such that DOB' and $D'OB$. Hence, by Theorem 2, $BO \equiv B'O$. Since $AB \equiv AB'$, $BO \equiv B'O$ and $AO \equiv AO$, therefore $\sphericalangle BAO \equiv \sphericalangle OAC$.

Corollary. If A, B, C are three non-collinear points and $AC \equiv AB$, there exists a point M such that BMC and $BM \equiv MC$ and AM is perpendicular to BC.

THEOREM 4 (A_{-2}).§ *If C is not on the line AB, there exists on the C side of the line AB (that side|| of the line AB which contains the point C) no point C' different from C such that $AC \equiv AC'$ and $BC \equiv BC'$.*

Proof. Suppose, on the contrary, that there exists on the C side of AB a point C' different from C and such that $AC \equiv AC'$ and $BC \equiv BC'$. Then by the Corollary to Theorem 3, there exists a point M such that CMC', $CM \equiv MC'$ and the line AM is perpendicular to the line CC'. Also there exists a point K such that CKC' and $CK \equiv KC'$, the line BK being perpendicular to the line CC'. By Theorem 23 of M. H. § 1, K is identical with M. But the lines MA and MB are distinct. This contradicts Theorem 1.

THEOREM 5. *If $\sphericalangle ABC \equiv \sphericalangle A'B'C'$ and D is in $\sphericalangle ABC$ and D' is in $\sphericalangle A'B'C'$ and $\sphericalangle ABD \equiv \sphericalangle A'B'D'$, then $\sphericalangle DBC \equiv \sphericalangle D'B'C'$.*

Proof. This Theorem follows from Theorem 18 of M. H. § 1 and A_{-2}.

---

* Cf. Theorem 7 and Definition 1 of M. H. § 1.

† Cf. Definition 11 of S. A.

‡ Cf. Theorem 6 of M. H. § 1.

§ See M. H. § 2, p. 492.

|| Cf. Definition 11 of S. A.

THEOREM 6. *If A is different from B, there exists a point M such that AMB and $AM \equiv MB$.*

Proof. By Theorem 10 of M. H. § 3, there exists a point C such that $\sphericalangle BAC$ is a right angle and a point D', in the plane ABC, such that $\sphericalangle ABD'$ is a right angle. Let D denote a point on the line BD' and on the C side of the line AB and such that $AC \equiv BD$. Let E be the point such that CAE and $AC \equiv AE$ and let F be the point such that DBF and $BD \equiv BF$. Then, by Axiom C_2 of M. H. and Theorem 5 of M. H. § 1, $AC \equiv BF$ and $AE \equiv BF$. By Axiom C_3 of M. H., $CE \equiv DF$. By Theorem 4 of M. H. § 1, $DF \equiv CE$, $BD \equiv AC$ and $BF \equiv AE$. Since CE is perpendicular to AB and $AC \equiv AE$, therefore $BC \equiv BE$ and $\sphericalangle CBA \equiv \sphericalangle EBA$. Since C and D are on the same side of the line AB, there is no point of AB between C and D. Similarly there is no point of AB between E and F. There is no point of AB between A and C or between A and E and, by Theorem 8 of M. H. § 3, the lines CE and DF have no point in common so that C is in $\sphericalangle ABD$ and E is in $\sphericalangle ABF$. Hence, by Theorem 5, $\sphericalangle CBD \equiv \sphericalangle EBF$. Hence $CD \equiv EF$ and $\sphericalangle CDB \equiv \sphericalangle EFB$. Therefore $CF \equiv DE$, $\sphericalangle ECF \equiv \sphericalangle FDE$, $\sphericalangle CED \equiv \sphericalangle DFC$ and $\sphericalangle CED \equiv \sphericalangle ECF$. The points C and F are on different sides of the line AB so that CF contains a point of the line AB and the lines CE and DF decompose the plane into three* regions, the segment AB containing all the points of the line AB in the region containing points between C and F. Hence there exists a point M such that AMB and CMF. Likewise there exists a point K such that CKF and DKE. Since $\sphericalangle CEK \equiv \sphericalangle ECK$, therefore, by Theorem 2, $CK \equiv EK$ and the line AK is perpendicular to the line CE. By Theorem 1, K is on the line AB. Hence K is M since K and M are common to the lines AB and CF. Since $\sphericalangle CEM \equiv \sphericalangle DFM$ and $\sphericalangle ECM \equiv \sphericalangle FDM$ and $CE \equiv DF$, therefore, by Theorem 1 of M. H. § 3, $CM \equiv DM$. But $\sphericalangle ACM \equiv \sphericalangle BDM$ and $AC \equiv BD$. Therefore $AM \equiv MB$.

* Cf. Theorem 27 of S. A.

THE UNIVERSITY OF TEXAS.

CLOSED POINT SETS ON A MANIFOLD.*

By S. Lefschetz.

I. Introduction.

1. The object of this paper is to investigate certain topological simultaneous invariants of a manifold M_n and a closed subset G on M_n. These invariants are integers analogous to the Betti numbers and include them as special cases. They satisfy certain relations (formulas (4), (5), (6), of No. 8) which include practically all duality relations now known in Combinatorial Analysis Situs.

When G is a complex C_k on M_n the invariants can be readily defined in terms of the complexes on C_k. For an arbitrary G however they may be defined in several modes depending upon the meaning attributed to the terms "cycle on G, complex on G or with boundary on G". The desideratum to be fulfilled is to preserve the properties obtained in the case of a C_k, and it is fulfilled by the sequences of "fractionary" cycles and complexes by means of which we define the cycles and complexes in question. Similar sequences of ordinary cycles were recently investigated by Vietoris† but appear to be inadequate for the purpose. However, in his paper (footnote 18) he refers to an invariant due to Brouwer which as we have shown (No. 27) coincides with our number $R_\mu(G)$ (generalized Betti number or connection index). Alexandroff‡ has also defined recently the Betti numbers of a closed set on an S_n as limits of those of suitable approximating complexes. It is not known as yet however if this mode of approach leads to invariants satisfying our relations except in a very special case,§ nor would it seem easy to prove. As a matter of fact the problem of the equivalence of the various modes of defining the invariants offers considerable interest.

Our general method consists in treating first C_k by isolating it within a neighborhood whose residual is a manifold with boundary. From a few

* Received December 1, 1927. See the two related preliminary Notes in Proc. Nat. Acad., vol. 13 (1927), pp. 614–622 and 805–807. We shall also mention repeatedly our two recent papers Trans. Amer. Math. Soc., vol. 28 (1926), pp. 1–49; vol. 29 (1927), pp. 429–462; and shall refer them as Trans. I, Trans. II. The parts actually vitally needed for the present paper are Trans. I, Part I, §§ 1, 2, 3, and Trans. II, Section II (mostly the duality theorems).

† Math. Ann., vol. 97 (1927), pp. 454–472.

‡ Comptes Rendus, vol. 184 (1927), pp. 317–319.

§ Alexandroff. Comptes Rendus, vol. 184 (1927), pp. 425–427.

of the simpler results obtained for the latter in Trans. II, we then derive the desired properties of C_k by means of a certain reduction theorem (Nos. 12, 13). In a somewhat similar fashion we pass from G to the general G.

One of our most suggestive results consists in two very general theorems on Kronecker-indices (intersections counted with multiplicities that may be positive or negative) for the cycles $M_n - G$ and the complexes with boundary on G. They are the direct extension of a noted theorem stated by Poincaré* and first proved independently by Veblen and Weyl,† and both reduce to it when G is a vacuous set. Once this proposition is established the duality formulas (4), (5) can be derived by essentially the same scheme, as indicated by Poincaré for the simplest case, where M_n has no boundary and G is vacuous.

2. **Notations and terminology.** The subscript in the symbol of a configuration shall always denote its dimensionality. It will frequently be omitted whenever there is no danger of confusion. We assume the reader familiar with what is meant by the terms *cell, complex, manifold, cycle* (Trans. I, p. 3; II, p. 437). The generic letters for these entities will be E, C, M, Γ, with variations when special types are involved. It is important to bear in mind that cells and complexes may be singular unless the contrary is explicitly stated. We recall that a singular E_k or C_k on A are subsets of A which are the continuous one way image (not necessarily one-one way image) of some ordinary E_k or C_k.

Concerning boundary relations we shall adopt instead of Poincaré's $\equiv$ the more convenient $\rightarrow$ recently introduced by Alexander.‡ Thus "the boundary of A is B" shall be expressed by $A \rightarrow B$ or $B \leftarrow A$. Let B be on D and let λB, $\lambda \neq 0$, be the boundary of some A on D. We shall express this relationship by $B \approx 0$, mod D. The symbol $\approx$ is the homology with division allowed or with zero divisors dropped, of our earlier writings.

Another convenient symbol is $\subset$ and its reverse to denote "is contained in" or "contains". Thus $B \subset A$, $A \supset B$ mean "B is a subset of A".

3. **Fractionary complexes and cycles.** We shall now introduce certain generalisations of great help in the sequel. We first extend the term *μ-complex* to cover a symbol $\frac{C_\mu}{q}$, where C_μ is an ordinary complex and q an arbitrary non zero integer. Let us call the usual type: *integral*

* Journal de l'Ecole Polytechnique, (2), vol. 1 (1895).

† Veblen, Trans. Amer. Math. Soc., vol. 25 (1923), pp. 540–550 (read before the Amer. Math. Soc. in 1920). Weyl. Revista Matem. (1923).

‡ Trans. Amer. Math. Soc., vol. 28 (1926), pp. 301–329.

complex, the new type *fractionary*—a distinction which we shall rarely need. By definition:

$$\frac{pC_\mu}{pq} = \frac{C_\mu}{q}; \qquad p \cdot \frac{C_\mu}{q} = \frac{pC_\mu}{q}; \qquad \frac{C_\mu}{q} \pm \frac{C'_\mu}{q'} = \frac{q'C_\mu \pm qC'_\mu}{qq'}.$$

A fractionary complex $\frac{C_\mu}{p}$ is called a *cycle*, whenever its numerator, the integral complex C_μ is a cycle. By "$\frac{C_\mu}{p}$ is on a configuration A or on $\frac{C_\nu}{q}$" we understand that the cells of C_μ are on A or on those of C_ν, it being unterstood that C_μ and C_ν are integral complexes. Henceforth all complexes and cycles are assumed fractionary unless otherwise stated, and we shall continue to apply the notations C_μ, Γ_μ, to them.

Regarding boundary relations and homologies we shall henceforth understand by $C_\mu \to C_{\mu-1}$ that there exists a $p \neq 0$ such that $pC_\mu \to pC_{\mu-1}$ in the ordinary sense: There is an integral complex pC_μ whose boundary is $pC_{\mu-1}$. When C_μ and $C_{\mu-1}$ are on the configuration A, we write corresponding to the boundary relation, $C_{\mu-1} \approx 0$ mod. A. $\to$, $\approx$ shall alone be used in the present paper to the exclusion of $\sim$ (ordinary homology). The effect of this limitation plus the introduction of fractionary cycles and complexes will be that in congruences and homologies coefficients may assume rational values and are not confined any more to integers. This is evidently no restriction in any investigation, such as the present one, wholly unaffected by the presence or absence of zero-divisors, and it is amply justified by the greater freedom acquired and by the simplification of many formulas.

It may also be said that fractionary cycles not essential when dealing with complexes alone, could not be dispensed with in the case of the general closed set. See in this connection No. 27.

Regarding the intersection theorems of Trans. I, II, we may apply them directly to the new types of complexes and cycles if we are careful to limit ourselves to the relation $\approx$.

4. **Topological invariants.** Let A be any point set such that we have assigned a meaning to the statement "a complex is on A". Consider the boundary relation

$$\sum \lambda_i C^i_{\mu+1} \to \sum \nu_j C^j_\mu \tag{1}$$

the complexes being all on A. The right hand side is a $\Gamma_\mu \approx 0$ mod. A. If Γ^h_μ, $h = 1, 2, \cdots, r$, are several cycles of A such that $\sum t_h \Gamma^h_\mu \approx 0$, mod. A, they are said to be *dependent* mod. A; when no such relation exists they are *independent*. The maximum number of independent μ-cycles

of A shall be denoted by $R_\mu(A)$. It is a topological invariant of A, provided that the *on* relation has been defined in an invariant manner for A, a remark which applies to our other invariants. It is certainly invariant when *on* means *carried by*. This will not always be the case in the sequel. Therefore with special definitions of the *on* relation it will be necessary to prove invariance directly.

Consider now a set B that need not intersect A but is such that, with the *on* relation for the complexes of A as to B is also defined. Let us neglect in (1) all complexes that are on B. Then in place of (1) we write

$$\sum \lambda_i' C_{\mu+1}^i \to \sum \nu_j' C_\mu^j, \text{ rel. } B. \tag{2}$$

The left side is then a cycle relatively to B (cycle rel. B), whenever the right is identically zero. An ordinary cycle is also a relative cycle but the reverse is not necessarily true. Since the right side of (1) is always an ordinary cycle (Trans. I, p. 5), the right side of (2) is always a relative cycle. The fact that it appears in a relation (2) is expressed by $\Gamma_\mu \approx 0$, mod. A, rel. B. The definition of dependence and independence is then as previously and the maximum number of independent relative cycles is denoted by $R_\mu(A; B)$. It is a topological simultaneous invariant of A and B, a statement to which must be applied the same reservations as for $R_\mu(A)$.

5. Further invariants arise when we consider the relations between the cycles of B and those of $A - B$. The carrying configuration A of all cycles is in general understood in any discussion and no provision is made to indicate it in the notation for invariants.* The latter are respectively maximum numbers of μ cycles:

On B, independent mod. B but ≈ 0 mod. A, or $r_\mu(B)$.

On $A - B$, independent mod. A of the cycles of B, or $S_\mu(A - B)$.

On A with no combination $\approx$ mod. A to a cycle on $B+$ a cycle on $A - B$, or $T_\mu(B)$.

Evidently $T_\mu(B) = T_\mu(A - B)$.

In regard to the various types of dependence, bounding or vacuous cycles are always to be taken in consideration. Thus if $\Gamma_\mu \subset A - B$, ≈ 0 mod. A, it is to be considered as $\approx$ a cycle on B mod. A — namely a vacuous cycle on B. It follows that $1 + R_\mu(A)$ μ-cycles of $A - B$ are always dependent mod. A, upon cycles of B and therefore $S_\mu(A-B) \leqq R_\mu(A)$. Hence *if* $R_\mu(A)$ *is finite the numbers* $S_\mu(A - B)$ *are all finite* and $\leqq R_\mu(A)$.

* In our Proceedings Notes, provision was made in the notation for the carrying configuration and for example instead of $r_\mu(B)$, we had $r_\mu(B; A)$ but the less precise notation adopted here has the advantage of being clearer end less cumbersome.

In particular when A is a complex as throughout our paper, $S_\mu(A-B)$ is always finite. Similarly $T_\mu(B) \leqq R_\mu(A)$ and is finite when the latter is finite, in particular for a complex.

6. A different class of invariants is obtained when all coefficients in boundary relations and homologies are reduced mod. p a fixed integer, which we assume prime in order to avoid complications, although this restriction is not essential. If the coefficients are fractions or if the cycles are fractionary the denominators must be assumed not divisible by p. This was done for $p=2$ by Veblen,* for p arbitrary by Alexander.† The new invariants thus obtained are denoted by $R_\mu(A;p)$, $R_\mu(A;B,p)$ etc.

The following generalisation suggests itself: Let $\mathfrak{C}$ be any additive property of the complexes $\subset A$. i. e. such that if C_μ and C'_μ possess it so does $\lambda C_\mu + \lambda' C'_\mu$, whatever the rational numbers λ, λ'. If we neglect throughout all complexes with the property $\mathfrak{C}$ we obtain what might be called boundary relations, homologies, mod. A rel. $\mathfrak{C}$, cycles rel. $\mathfrak{C}$. The maximum number of independent μ-cycles of this nature is a number that may be designated by $R_\mu(A;\mathfrak{C})$. Thus we get the invariants $R_\mu(A;B)$ when $\mathfrak{C}$ means: $C_\mu \subset B$, the invariants $R_\mu(A;p)$ when $\mathfrak{C}$ means: C_μ is p times some other complex. It would appear to be highly interesting to find out if the total class of topological invariants thus obtainable, say for a complex, includes any that are essentially different from those of the present paper, i. e. corresponding to the two special cases of $\mathfrak{C}$ just mentioned or to their combinations.

7. **Generalized Euler-Poincaré formula.** Let $A = C_n$, a non singular complex, and $B = C_k$, a subcomplex of C_n. Let α_i denote the number of i-cells of $C_n - C_k$. Then we have the generalized Euler-Poincaré relation,

$$\sum(-1)^i \alpha_i = \sum(-1)^i R_i(C_n; C_k, p). \tag{3}$$

The ordinary Euler-Poincaré relation corresponds to C_k vacuous, $p=1$, i. e. involves the numbers $R_i(C_n)$.‡ For C_k vacuous, $p=2$, the formula has been derived by Veblen and Alexander,§ and for $p>2$ by Alexander.‖ Their proofs apply to (3) directly and need not be repeated here.

8. **Duality relations.** Let M_n be the set of points on a complex C_n, and G a closed set $\subset M_n$ which includes all cells of C_n that do not verify

* Annals of Math., (2), vol. 14 (1912), pp. 86–94; see also Veblen-Alexander, ibid pp. 163–178; Veblen, Colloquium Lectures on Analysis Situs, Chs. 1, 2, 3.

† In the Transactions paper already quoted.

‡ It was first proved by Poincaré in the paper already quoted.

§ In the joint paper just quoted. See also Veblen, Coll. Lect. p. 78.

‖ In the paper of the Trans. just quoted.

the conditions for an n-dimensional manifold with boundary (Trans. II p. 437). M_n may be considered as an n-dimensional manifold not necessarily connected with a boundary $F_{n-1} \subset G$ or with a subcomplex of singularities also $\subset G$. Strictly speaking the notation "M_n" should not be applied to it, but if one keeps in mind the more general meaning adopted here no harm will follow. Orientability can be defined for C_n as for any manifold, if one applies the usual scheme only to $(n-1)$-cells incident with two n-cells. Unless otherwise stated we shall assume M_n orientable. C_n is the *defining* complex of M_n, and in order to have no doubt as to what we shall do, we agree that given G, any other defining C_n to be considered must have its associated subcomplex of singularities likewise $\subset G$. In order to bring out clearly the degree of generality implied here we may point out that our treatment applies to complexes whose non boundary cells alone satisfy the manifold conditions. For example they may be manifolds in the sense of Veblen: Manifolds without boundary with some n-cells of a defining C_n removed. The fact is that we shall actually operate constantly not on M_n but on certain manifolds $M_n - G$ and of the more restricted types of Trans. II, p. 437. The actual behavior of M_n at the boundary is therefore an indifferent matter.

This being understood, the simultaneous invariants of M_n and G satisfy the following two basic duality relations

$$S_\mu(M_n - G) = S_{n-\mu}(M_n - G) \tag{4}$$

$$R_\mu(M_n - G) = R_{n-\mu}(M_n; G). \tag{5}$$

In addition we have the formula

$$R_\mu(M_n; G) = r_{\mu-1}(G) + S_\mu(M_n - G) + T_\mu(G). \tag{6}$$

One of our chief objects will be to prove these three relations. From them we derive easily

$$R_\mu(M_n - G) - S_\mu(M_n - G) = r_{n-\mu-1}(G) + T_{n-\mu}(G) \tag{7}$$

which is essentially (4) of our first Proceedings Note, where the left hand side has been designated by s_μ: It is the maximum number of distinct μ-cycles of $M_n - G$ dependent upon cycles of G mod. M_n.

When G is vacuous, which requires that M_n be without boundary, both (4) and (5) reduce to Poincare's noted duality relation

$$R_\mu(M_n) = R_{n-\mu}(M_n).$$

On the other hand when M_n is an n-sphere H_n, (7) embodies the generalisation of Alexander's duality relation between the invariants of

a $C_k \subset H_n$ and those of $H_n - C_k$ (formula (9) below). Thus *our group of formulas is, so to speak, the point of convergence of two apparently only distantly related questions of topology.*

By merely inserting the integer p in the proper place in each invariant we have new formulas for the cycles mod. p which are also true and derived exactly like the others. The same is also true when taking the invariants for cycles rel. H, a closed set having no points in common with G. When H and G have common points the problem becomes more complicated and will not be taken up here.

An important observation due to Veblen and Alexander is that the invariants mod. 2 can also be considered for a non orientable M_n. In a sense they embody results of a strictly point-set theoretic nature, with all questions of orientation left out.

9. In our treatment we shall first take up at length the case of a $C_k \subset M_n$, then the general G. But before doing so we shall give bibliographical indications.

Important cases of (4), (5) occur in the literature: When G is a vacuous set, both reduce to the noted duality relation of Poincaré.* The similar relation for cycles mod. 2 has been derived by Veblen and Alexander,† for cycles mod. any p by Alexander.‡ When $G = F_{n-1}$ boundary of M_n, (4) and (5) reduce to certain duality relations due to the author.§ For $G = C_k$ a subcomplex of the defining C_n of M_n they were first derived in our first Proceedings Note.

When $M_n = H_n$ a hypersphere, $S_\mu(H_n - G) = 0$ for every μ; $T_\mu(G) = 0$ for $\mu \neq n$, $T_n(G) = 1$, since the only cycles not ≈ 0 carried by H_n are a point and H_n itself. This is likewise true for the invariants mod. p. Let $T_\mu = 0$ for $\mu \neq n$, $T_n = 1$. Then (5) together with (6) and the remark that in the present case $r_\mu(G) = R_\mu(G)$, give:

$$R_\mu(H_n - G) = R_{n-\mu-1}(G) + T_{n-\mu}. \tag{8}$$

The same formula for $G = C_k$ and cycles mod. 2

$$R_\mu(H_n - C_k; 2) = R_{n-\mu-1}(C_k; 2) + T_{n-\mu} \tag{9}$$

was first derived by Alexander.‖ Its importance is manifest when it is remembered that by means of it he showed how to derive the generalized

* Palermo Rendiconti, vol. 13 (1899).

† In their joint paper, also Veblen, Coll. Lect., p. 91.

‡ In his paper of the 1926 Transactions.

§ Trans. II, pp. 449, 450, first and second duality theorems. In the statement of the first, $n - \mu$ must be replaced by $n - 1 - \mu$ as the discussion clearly shows.

‖ Trans. Amer. Math. Soc., vol. 23 (1922), pp. 333–349.

Jordan-Brouwer theorem in a few lines. A still different method has been recently applied by Alexandroff* to the derivation of (9) for $\mu = 0$ and the most general G. The numbers $R_0(H_n - G; p)$ are in fact all equal and their common value $R_0(H_n - G)$ is the number of disconnected domains into which H_n is separated by G. This invariant is therefore the simplest and at the same time the most important simultaneous topological invariant of H_n and G.

Another important result implied by (9), and the similar formula with p in place of two, is that they give expressions for the topological invariants $R_\mu(G; p)$ of G in terms of those of its residual set $H_n - G$,—a fundamental remark due to Alexander.

II. The general C_k on M_n.

10. From now on we are considering complexes and sets on an M_n which we assume orientable unless otherwise stated. The manifold possesses a certain defining complex C_n. On M_n we shall have occasion to consider intersections of complexes but the situation will always be as in Trans. I: The boundaries of intersecting complexes will not meet. Let C_μ, C_ν be two such intersecting complexes. We shall usually be concerned with the intersection cycle $C_\mu \cdot C_\nu$, or Kronecker index $(C_\mu \cdot C_\nu)$ when $\mu + \nu = n$. In view of what has been shown loc. cit. C_μ may be replaced by $\bar{C}_\mu$, $(C_\mu^0 + C_\mu'$ of Trans. I) with the same boundary and polyhedral down to cells as close as we please to that boundary. There is a similar $\bar{C}_\nu$ for the other complex and the two approximations may be so chosen that the intersecting cells be all polyhedral and in the most general position relatively to one another on C_n. Whenever intersecting complexes occur in the sequel all complexes will be assumed replaced by the approximations in general position relatively to one another.

11. Let then C_k be a fixed complex on M_n. As a matter of fact this particular complex might be replaced throughout by the homeomorph of any polyhedral configuration in some Euclidean space, composed of a finite number of faces of dimension up to k and constituting a closed point set. If M_n *has a boundary it must be assumed* $\subset C_k$. The latter is non singular but not necessarily polyhedral, i. e., it has the same degree of generality as, say a Jordan curve. Our first objective is an important reduction theorem for complexes very near C_k which will play an important

* Comptes-Rendus, vol. 184 (1927), pp. 425–427; earlier for $n = 2$, G an arbitrary closed set Brouwer, Math. Ann., vol. 72 (1912), p. 422; Alexandroff from another point of view, Math. Ann., vol. 96 (1926), pp. 527–534. As a matter of fact as he pointed out to me his $R_{n-1}(G; 2)$ is not defined like ours. But (9) being proved for both it follows that they are equal, their common value being $R_0(H_n - G) - 1$. Further references at the end of this paper.

part in the discussion. In its proof we need the following property which is best taken up separately: *Let H_i be an i-sphere (singular cell boundary) on C_k. For every ξ sufficiently small there is an η (tending to zero with ξ) such that when the diameter of $H_i < \eta$, $C_k \subset$ an E_{i+1} (in general singular) bounded by H_i and of diameter $< \xi$.* To show it subdivide C_k into C_k' with cells of diameter $< \frac{1}{2}\xi$. Then there exists an η tending to zero with ξ, and such that any subset of C_k' of diameter $< \eta$, $\subset$ a set of cells of C_k' having a common vertex (star of cells with that vertex for center). If H_i is that set and A the vertex all that is necessary is to join it to all points of H_i by one-cells rectilinear with respect to C_k and its definition of straightness. Their locus is the required E_{i+1}.

12. **Reduction theorem***. *Let N^ε denote the set of all points of M_n at a distance $< \varepsilon$ from C_k. To each $\varepsilon > 0$ corresponds $\varepsilon' > 0$, tending to zero with ε, and such that if $C_\mu \subset N^\varepsilon$, $C_\mu \to \Gamma_{\mu-1} \subset C_k$, then also $C_\mu \approx C_\mu'$ mod. $N^{\varepsilon'}$, $C_\mu' \subset C_k$.*

In other words a cycle rel. C_k which is sufficiently near C_k is ≈ 0 mod. $N^{\varepsilon'}$ rel. C_k. (If the cycles and complexes were all integral we could replace $\approx$ by $\sim$.)

Subdivide C_μ into cells of diameter $< \varepsilon$, and continue to call the new complex C_μ. This will lead to no confusion since it is $\approx$ the initial one mod. N^ε. Let then $A^1, \cdots, A^\alpha$ be the vertices of C_μ. Associate with each A^i the point B^i nearest to it, on C_k, or one of the set of points at the least distance if there is more than one. Let A^iA^j be a one-cell of C_μ. The points B^i, B^j are at distance $< 3\varepsilon$ from one another. Hence by No. 11 if ε is sufficiently small there will be a one-cell B^iB^j on C_k and furthermore its diameter can be made as small as desired. Similarly to any E_2 say $A^iA^jA^h$ of C_μ can be made to correspond an E_2; $B^iB^jB^h$ on C_k; etc. The only condition is that $\varepsilon <$ a certain number which depends upon several ξ's such as in the preceding number, but not upon C_μ. Furthermore we assume that when any cell of C_μ is on C_k the corresponding B cell coincides with it.

Apply now the result of No. 11 to M_n itself: We can find one-cells A^iB^i whose diameter can be made as small as we wish by taking ε sufficiently small. If A^iA^j is an E_1 of C_μ, $A^iA^jB^jB^i$ is an H_1 whose diameter depends upon ε and tending to zero with ε. Hence for ε sufficiently small there is an E_2 of which it is the boundary, etc. Proceeding thus we shall have a $C_{\mu+1}$ whose cells are of a diameter $< \varepsilon'$ tending to zero with ε, hence on $N^{\varepsilon'}$. We assume again that where an E_h of C_μ is on C_k, the

*See our Proceedings Notes whose proof of the theorem we have greatly simplified; the present treatment is very close to Alexander's, Trans. Amer. Math. Soc., vol. 16 (1915), pp. 148–154. In a similar order of ideas see Veblen, Coll. Lect., pp. 94, 119.

corresponding E_{h+1} of $C_{\mu+1}$ is a singular cell coincident with E_h. The incidence relations between the $\mu+1$ cells of $C_{\mu+1}$ and its μ cells of type $A^1 \cdots A^\mu B^1 \cdots B^\mu$ are the same as between the μ and $\mu-1$ cells of C_μ. Let us then orient each $E_{\mu+1}$ of $C_{\mu+1}$ positively relatively to the associated E_μ of C_μ. Since in C_μ the boundary cells of the μ-cells not on C_k cancel one another the same will hold for the boundary cells of $C_{\mu+1}$ not on C_k and not cells of C_μ. In other words $C_{\mu+1} \to C_\mu +$ a complex on C_k, which proves the theorem.

13. **Polyhedral neighborhood of C_k.*** In order to be able to apply the results of Trans. II we must replace N^ε by a suitable polyhedral neighborhood. To obtain it we first subdivide C_n into C_n^1 with cells of diameter $< \frac{1}{2}\varepsilon$ and take the subcomplex C_n' of C_n^1 made up of all cells which carry a point of C_k or have one on their boundary, plus the boundaries of all such cells. The cells of C_n^1 abutting on C_n' constitute together with their boundaries and C_n' a new C_n'' also a subcomplex of C_n^1. The cells of C_n'' not on C_n', nor abutting on it constitute a C_{n-1}. Let PQ be a segment of $C'' - C'$ whose end points are respectively on C' and C_{n-1}, and let R be a point dividing it in any fixed ratio $\lambda < 1$. As PQ describes any particular E_h of $C'' - C'$, R describes an E_{h-1}, polyhedral, convex and flat, and the locus of all these cells is a complex Φ_{n-1}. The set of all closed segments PR plus C_n' is an n-complex N *which is the required polyhedral neighborhood of* C_k. It is a *closed* neighborhood, the corresponding open neighborhood being $N - \Phi$. Both N and Φ have all their points within a distance ε of C_k and as near to it as we please. ε is the *width* of N. Moreover, the points of Φ are all at a distance > 0 from C_k, and C_k is wholly interior to N.

Of equal importance with N is the n-complex $K_n = M_n - N + \Phi$, likewise polyhedral and with convex, flat, cells. Its points are also all at a distance > 0 from C_k. We need the following properties:

I. *In the reduction theorem of No.* 12, N^ε, $N^{\varepsilon'}$ *may be replaced by polyhedral neighborhoods* N *and* $N' \supset N$ *of widths* ε, ε'.

II. K_n *is an* M_n. The definition of manifolds here used is strictly that of Trans. II, p. 437. It implies that the boundary Φ of K is also a manifold if K itself is one. To prove the proposition we need only to consider the behavior of the cells of Φ relatively to K. Any cell E_j' of Φ is on a certain E_{j+1} of C_n'. Every E_h', $h > j+1$, of $K - \Phi$ incident with E_j' is on an E_h of C_n' incident with E_{j+1} and the incidence relations between corresponding cells of the two sets $\{E_h\}$, $\{E_h'\}$ are the same. Hence since

* The neighborhood here given was considered at length in our first Proceedings Note. It was also obtained independently and by the same construction by Alexandroff, Math. Ann., vol. 98.

C_n' defines a manifold with E_{j+1} for interior cell, the two sets have the structure of an H_{n-j-1}. Hence, E_{j+1}' being the subcell of K on E_{j+1}, (the part of E_{j+1} on K), the set $E_{j+1}' + \{E_h'\}$ has the structure of an E_{n-j-1} plus its boundary, every E_i of the set corresponding to an E_{i-j-1} of a subdivision of (E_{n-j-1} + its boundary) into cells. Referring to our definition of an M_n with boundary, this proves that K is such a manifold with Φ for boundary.

In the same manner it may be shown that N is an M_n, but that is immaterial for the sequel.

Remark. It will frequently happen that some of our statements referring to a neighborhood N of C_k will hold only when its width ε is chosen sufficiently small. Such a choice, always possible wherever it will be implied, shall be automatically assumed without anything further being said about it.

III. Theorems on Kronecker indices.

14. From now on in this section we shall use Δ for cycles rel. C_k and continue to use Γ as the generic letter for ordinary cycles.

In the theorems to be stated presently there appears the Kronecker index of two complexes. We recall that the index $(C_h \cdot C_{n-h})$ of C_h, $C_{n-h} \subset M_n$ is the number of their intersections each counted with a certain multiplicity that may be positive or negative (Trans. I, pp. 6, 12). In the proof we shall make repeated application of the following proposition also true for fractionary cycles and complexes (Trans. I, p. 20, Th. IV): *If* $C_{\mu+1} \to \Gamma_\mu$ *and if* $C_{n-\mu}$ *is a complex whose boundary does not meet* $C_{\mu+1}$, *then* $(\Gamma_\mu \cdot C_{n-\mu}) = 0$. *In particular this holds when* $C_{n-\mu}$ *is a cycle.* To illustrate this basic theorem by a very simple example, let γ be the boundary of a spherical region R, δ a spherical arc whose end points are not on $R + \gamma$. Then the theorem merely states the obvious fact that if a point describes δ in a definite direction it enters R as often as it leaves it.

15. Theorems. I. *In order that* Γ_μ, *a fixed cycle of* $M_n - C_k$, ≈ 0 *mod.* $M_n - C_k$ *it is necessary and sufficient that* $(\Gamma_\mu \cdot \Delta_{n-\mu}) = 0$ whatever $\Delta_{n-\mu}$.

II. *Conversely in order that* $\Delta_{n-\mu} \approx 0$ mod. M_n rel. C_k, *it is necessary and sufficient that the index be zero whatever* Γ_μ.

When C_k is vacuous I and II reduce to the theorem of Veblen and Weyl mentioned in the Introduction. From the theorem recalled in No. 14 follows immediately that the conditions in I, II are necessary. For when $M_n - C_k \supset C_{\mu+1} \to \Gamma_\mu$, where $C_{\mu+1}$ does not meet the boundary of Δ which is on C_k, necessarily $(\Gamma \cdot \Delta) = 0$. Next when $\Delta \approx 0$ mod. M_n rel. C_k, $C_k \supset C_{n-\mu}$ and $M_n \supset C_{n-\mu+1} \to \Delta_{n-\mu} + C_{n-\mu}$. Since $C_{n-\mu}$ does not meet Γ_μ, $(\Gamma_\mu \cdot C_{n-\mu}) = 0$, hence $(\Gamma_\mu \cdot \Delta_{n-\mu} + C_{n-\mu}) = 0 = (\Gamma \cdot \Delta)$.

Therefore we merely have to prove the sufficiency of the conditions.

16. *Consider first the case where* $C_k = F_{n-1}$, *boundary of* M_n, a manifold in the strict sense of Trans. II p. 437. As p. 439 of that paper we take a copy $\bar{M}_n$ of M_n and match the two along associated boundary points so as to have a manifold without boundary $V_n = M_n - \bar{M}_n$. The image of any particular element of $\bar{M}_n$ on M_n will be denoted by the same letter barred.

We have shown (Trans. II No. 19 Th. I), that every cycle of V_n is homologous to a sum of: (a) A cycle on M_n, which may be assumed interior to it (ibid. No. 23 Coroll. I), and therefore is a Γ in our notation. (b) A skew-symmetric cycle, i. e., of type $\Delta - \bar{\Delta}$ in our present notation. Hence on V_n every $n-\mu$ cycle $\Gamma'_{n-\mu} \approx \Gamma_{n-\mu} + \Delta_{n-\mu} - \bar{\Delta}_{n-\mu}$, a very simple result that may be proved directly with ease.

17. Let now Γ_μ behave as in I. Then $(\Gamma_\mu \cdot \Gamma'_{n-\mu}) \approx (\Gamma_\mu \cdot \Delta_{n-\mu}) = 0$ by assumption, and $(\Gamma_\mu \cdot \bar{\Delta}_{n-\mu}) = 0$ because the two complexes do not meet, the first being on $M_n - F$, the second on $\bar{M}_n$; it follows $(\Gamma_\mu \cdot \Gamma'_{n-\mu}) = 0$ whatever Γ' and by the theorem of Veblen and Weyl, $\Gamma_\mu \approx 0$ mod. V_n. Hence (loc. cit. No. 18, Th. IV) $\Gamma_\mu \approx 0$ mod. M_n which completes the proof of I in the present instance.

18. Let now $\Delta_{n-\mu}$ behave in accordance with II. Let $\delta_{n-\mu} = \Delta - \bar{\Delta}$. Then, whatever the skew-symmetric cycle δ_μ, $(\delta_\mu \cdot \delta_{n-\mu}) = 0$ (Trans. II, No. 24, Th. II). Furthermore since every cycle of $M_n \approx$ mod. M_n to an interior cycle, $(\Gamma_\mu \cdot \delta_{n-\mu}) = (\Gamma_\mu \cdot \Delta_{n-\mu}) = 0$ by assumption whatever Γ_μ, on M_n. By the theorem on the cycles of V_n recalled above, this holds then whatever Γ_μ on V_n. Therefore $\delta_{n-\mu} \approx 0$ mod. V_n. Hence $V_n \supset C_{n-\mu+1} \to \delta_{n-\mu}$. Since in Δ all $n-\mu$ cells on F_{n-1} may be neglected (as ≈ 0 mod. M_n, rel. F_{n-1}), we may assume that it has none. Furthermore we can manifestly take all complexes polyhedral and C in the form $C' + C''$, $C' \subset M_n$, $C'' \subset \bar{M}_n$. The cells of Δ not on F_{n-1} can only belong to the boundary of C'. Hence $C' \to \Delta$, rel. F_{n-1}, and $\Delta \approx 0$ mod. M_n, rel. F_{n-1} which completes the proof of II for the case considered.

19. We now take up the case *where* C_k *is our former general complex on* M_n and reduce it to the case just considered. It must be remembered that when M_n has a boundary F_{n-1}, always $F_{n-1} \subset C_k$.

Let then Γ_μ be such that $(\Gamma_\mu \cdot \Delta_{n-\mu}) = 0$ for every Δ. Consider two polyhedral neighborhoods N, N^1 associated after the manner of No. 13 ($N^1 \supset N$, but their widths tend simultaneously to zero), and so chosen that $\Gamma_\mu \subset M_n - N^1$. Let $C^1_{n-\mu}$ be an arbitrary complex of K with its boundary $\Gamma_{n-\mu-1}$ on Φ. From No. 13 property I follows that N^1 carries a $C^2_{n-\mu} \to -\Gamma_{n-\mu-1} +$ a cycle on C_k. Hence $C^1 + C^2$ is a $\Delta_{n-\mu}$ so that $(\Gamma \cdot C^1 + C^2) = 0$. But since $C^2 \subset N^1$ and $\Gamma \subset M_n - N^1$, C^2 and Γ do not meet. Hence $(\Gamma \cdot C^2) = 0$, and finally $(\Gamma \cdot C^1) = 0$. But C^1 is

an arbitrary cycle of K, rel. Φ. Hence by the part of the theorem already proved, $\Gamma \approx 0$ mod. K, hence mod. $M_n - C_k$ which completes the proof of I.

Regarding II, let $\Delta_{n-\mu}$ be such that $(\Gamma \cdot \Delta) = 0$ whatever Γ. We have $\Delta_{n-\mu} = C^1_{n-\mu} + C^2_{n-\mu}$; $C_1 \subset K$; $C_2 \subset N$. An arbitrary Γ_μ on $K - \Phi$ does not meet C^1, hence $(\Gamma \cdot C^2) = 0 = (\Gamma \cdot \Delta) = (\Gamma \cdot C^1 + C^2) = (\Gamma \cdot C^1)$. Therefore by the part of II already proved $C^1 \approx 0$, mod. K, rel. Φ. This means that $C^1 \approx C^3 \subset \Phi$, mod. K hence mod. M_n. Therefore $\Delta \approx C^2 + C^3$ mod. M_n; $C^3 \subset N$. The boundary of $C^2 + C^3$ is that of Δ, hence it is on C_k. Therefore (No. 12, property I of No. 13), $C^2 + C^3 \approx$ a complex on C_k mod. N^1, hence mod. M_n; therefore finally $\Delta \approx 0$, mod. M_n, rel. C_k which completes the proof of II.

IV. Proof of the relations between the invariants of C_k.

20. **Proof of (4).** We have seen that the numbers $S_\mu(M_n - C_k)$ are all finite (No. 5). Let Γ^j_μ, $j = 1, 2, \cdots, S_\mu$, be a maximal set of cycles of $M_n - C_k$, independent mod. M_n rel. C_k, consider $\|(\Gamma^j_\mu \cdot \Gamma^h_{n-\mu})\|$ and assume $S_\mu > S_{n-\mu}$. There exists then a linear combination of the rows composed wholly of zeros, or a non-trivial $\Gamma_\mu = \sum \lambda_j \Gamma^j_\mu$ such that every $(\Gamma_\mu \cdot \Gamma^h_{n-\mu}) = 0$. Since every $\Gamma_{n-\mu} \approx \sum \nu_h \Gamma^h_{n-\mu}$ mod. M_n rel. C_k, $(\Gamma_\mu \cdot \Gamma_{n-\mu}) = \sum \nu_h (\Gamma_\mu \cdot \Gamma^h_{n-\mu}) = 0$, and by Th. II on Kronecker-indices, $\Gamma_\mu = \sum'_j \lambda_j \Gamma^j_\mu \approx 0$, mod. M_n rel. C_k, contrary to assumption.—Hence $S_\mu \leqq S_{n-\mu}$. Similarly of course $S_{n-\mu} \leqq S_\mu$. Hence the two numbers are equal or

$$S_\mu(M_n - C_k) = S_{n-\mu}(M_n - C_k) \tag{10}$$

which is (4) for $G = C_k$.

21. **Proof of (5).** Let Γ^h_μ, $h = 1, 2, \cdots, R_\mu(M_n - C_k)$; $\Delta^j_{n-\mu}$, $j = 1, 2, \cdots$ $\cdots, R_{n-\mu}(M_n; C_k)$ be maximal sets of independent cycles of $M_n - C_k$, and of M_n rel. C_k. Consider the associated matrix $\|(\Gamma^h \cdot \Delta^j)\|$ and let $R_\mu(M_n - C_k) > R_{n-\mu}(M_n; C_k)$. As a consequence there exists a linear combination of the rows of the matrix composed wholly of zeros. Hence there is a non trivial linear combination Γ_μ of the cycles Γ^h_μ, such that every $(\Gamma \cdot \Delta^j) = 0$. Let $\Delta_{n-\mu}$ be an arbitrary cycle rel. C_k. Then $\Delta \approx \sum \lambda_j \Delta^j$, mod. M_n rel. C_k. Hence (Th. II on Kronecker indices), $(\Gamma \cdot \Delta) = \sum \lambda_j (\Gamma \cdot \Delta^j) = 0$. Hence (Th. I), $\Gamma \approx 0$ mod. $M_n - C_k$. This implies dependence of the cycles Γ^h mod. $M_n - C_k$ contrary to assumption. The inequality assumed is therefore false. That its inverse is likewise false is proved in exactly the same way with the parts of the Γ's and Δ's as well as homologies mod. $M_n - C_k$ and mod. M_n rel. C_k interchanged. Hence

$$R_\mu(M_n - C_k) = R_{n-\mu}(M_n; C_k) \tag{11}$$

which is (5) for $G = C_k$.

22. **Proof of (6).** In our case it reads

$$R_\mu(M_n; C_k) = r_{\mu-1}(C_k) + S_\mu(M_n - C_k) + T_\mu(C_k). \tag{12}$$

We can immediately single out relative cycles in number equal to the right hand side: It consists of:

(a) The set Γ^j_μ, $j = 1, 2, \cdots, S_\mu$ of Nr. 20.

(b) A set D^j_μ, $j = 1, 2, \cdots, T_\mu(G)$, composed of cycles independent of the Γ's mod. M_n rel. C_k.

(c) In addition a set of relative cycles Δ^j_μ, $j = 1, 2, \cdots, r_{\mu-1}$, obtained thus: There is on C_k a maximal set of cycles $\gamma^j_{\mu-1}$, $j = 1, 2, \cdots, r_{\mu-1}$, independent mod. C_k but each ≈ 0 mod. M_n. There exists then a $\Delta^j_\mu \to \gamma^j_{\mu-1}$, and these Δ's are the relative cycles in question.

To prove (12) we must show that the relative cycles of the three sets are independent mod. M_n rel. C_k, and that any other depends upon them.

Assume then first a relation

$$\sum \lambda_j \Gamma^j + \sum \nu_j D^j + \sum \varrho_j \Delta^j \approx 0, \text{ mod. } M_n \text{ rel. } C_k. \tag{13}$$

Choose again K such that the Γ's are all on $K - \Phi$. By taking intersections with Φ and applying Th. I of Trans. I, p. 12, we derive from (13)

$$\sum \nu_j D^j \cdot \Phi + \sum \varrho_j \Delta^j \cdot \Phi \approx 0, \text{ mod. } \Phi. \tag{14}$$

Now $D^j \cdot \Phi \leftarrow$ the part of D^j on N and $\Delta^j \cdot \Phi - \gamma^j \leftarrow$ the part of Δ^i on N; hence they are ≈ 0 mod. N. From this and (14) follows $\sum \varrho_j \gamma^j \approx 0$ mod. N. Hence $N \supset C_\mu \to \sum \varrho_j \gamma_j$. By the reduction theorem (Nos. 12, 13) C_k carries a complex $\approx C_\mu$, hence $\to \sum \varrho_j \gamma^j$ like C_μ. Therefore $\sum \varrho_j \gamma^j \approx 0$ mod. C_k contrary to our assumption as to the γ's. It follows that every $\varrho = 0$. We conclude then that (14) reduces to

$$D \cdot \Phi \approx 0, \text{ mod. } \Phi, \qquad D = \sum \nu_j D^j.$$

Hence $\Phi \supset C_\mu \to D \cdot \Phi$. Let $D = D' + D''$, $D' \subset K$, $D'' \subset N$. We have $D' \to D \cdot \Phi$, $D'' \to -D \cdot \Phi$. Also $D = (D' - C_\mu) + (D'' + C_\mu)$. Each parenthesis represents a cycle, the first $\subset K$, hence dependent upon the cycles Γ^j mod. $M_n - C_k$, the second $\subset N$, hence $\approx$ a cycle on C_k mod. M_n (Nos. 12, 13). Therefore $D = \sum \nu_j D^j$ turns out to be dependent upon the Γ's mod. M_n rel. C_k — contrary to the assumption of mutual independence of the sets Γ^j, D^j. It follows that every $\nu = 0$. But (13) is then reduced to a relation on the Γ's alone, which we have assumed that they do not satisfy, hence every λ is zero. Thus (13) cannot hold and the relative cycles considered are independent.

Let now Δ_μ be any relative cycle whatever, $\to \gamma_{\mu-1} \subset C_k$. In view of the maximal property of the cycles γ^j, $\gamma \approx \sum \varrho_j \gamma^j$, mod. C_k. Hence $C_k \supset C_\mu \to \gamma - \sum \varrho_j \gamma^j$. Therefore $\Delta - \sum \varrho_j \Delta^j - C_\mu$ is a cycle and from the definition of the Γ's and D's follows that it depends upon them and upon cycles on C_k itself. Hence

$$\Delta \approx \sum \varrho_j \Delta^j + \sum \lambda_j \Gamma^j + \sum \nu_j D^j, \text{ mod. } M_n, \text{ rel. } C_k.$$

Thus the Δ's, Γ's, D's do constitute a maximal set of independent relative cycles from which (12) follows.

V. The closed set G on M_n.

23. **Complexes and cycles on G.** There is no a priori reason for not making the definition of the topological invariants rest upon complexes whose cells are on G in the ordinary sense and indeed there is thus obtained a definite class of topological invariants. In general however it is not possible to prove for such complexes a reduction theorem like that of No. 12 which has played a basic part in our treatment of the case of C_k. The related invariants will therefore fail to satisfy our basic relations (4), (5), (6) except in special cases. It points to a lack of connection between the cycles and complexes on G in the ordinary sense and those in a suitable neighborhood of G.

It will be profitable to examine in this regard a simple example due to Alexandroff.* G consists of the following subset of the Euclidean plane H_2: $y = \sin \frac{1}{x}$, $0 < x \leqq 1$, plus the segment $-1 \leqq y \leqq +1$, $x = 0$, plus a polygonal line joining the points (0, 1), (1, sin 1) and not carrying any of the points already assigned to the set. The total set, as is well known, is not locally connected from which readily follows that it carries no linear cycle not deformable into a point. Hence by means of *arcs on* G in the strict sense one can only obtain the value $R_1(G) = 0$. On the other hand any neighborhood of width ε of G carries a Γ_1 which is ≈ 0 mod. a neighborhood of width ε' that *does not* tend to zero with ε. From this and in the light of what we shall say below (No. 28) it follows that the number $R_1(G)$ to be defined presently is equal to *one*. Since our basic formulas will be proved for our topological invariants we can then conclude from (9), $R_0(H_2 - G) = 2$. This shows that G, like the Jordan curve, decomposes the plane into *two* connected regions—a conclusion which could not be deduced from any known relation for invariants of G obtained when the *on* relation is taken in its strict sense.

* Math. Ann., vol. 96 (1926), pp. 489-511.

The key to what follows is now plain: We must define the notion "complex on G" in such a manner as to make the basic relations (4), (5), (6) hold. This condition, as we shall show, is fulfilled by our definitions.

24. Let us again subdivide C_n into C_n' with cells of diameter $< \varepsilon$, then remove all cells carrying a point of G in their interior or on their boundary. The totality of these cells plus their boundaries is now the closed polyhedral neighborhood to be designated by N. The common boundary of N and $M_n - N$ is an $n-1$ complex Φ, and $M_n - N + \Phi = K$ is an n-complex. These complexes will play the same part as those of same name considered previously. We could carry the construction farther and obtain manifolds K, N, Φ, but by taking advantage of the results already acquired we can dispense with the extra step.

Consider now a sequence $\{N^i\}$ of our neighborhoods of widths tending to zero with $\frac{1}{i}$, and such that $N^i \supset N^{i+1}$ for every i, but otherwise unrestricted. Thus they do not even have to be constructed by means of the same defining C_n of M_n. One of them may for example be polyhedral in regard to a given C_n, another not, etc.

We now define a *μ-cycle on G* as a sequence $\gamma_\mu = \{\Gamma_\mu^i\}$ such that $\Gamma^i \subset N^i$ while for every p there is a q with the property that when $i, j \geq q$, $\Gamma^i \approx \Gamma^j \bmod. N^p$.

We shall free the definition from the particular $\{N^i\}$ considered by agreeing that if cycles $\{\Gamma_\mu^i\}$ defined by means of $\{N^i\}$ and $\{\Gamma'^i_\mu\}$ defined by means of $\{N'^i\}$ have a common infinite subsequence of elements Γ they define one and the same γ_μ. Each sequence $\{\Gamma_\mu^i\}$ may then be considered as a *representation* of γ_μ relatively to $\{N^i\}$. The cycle appears thus as a class of representations such that any two are extremes of a chain of representations with every pair of consecutive terms related as above, the number of terms being of course finite.

From a representation $\{\Gamma_\mu^i\}$ relatively to $\{N^i\}$ there can always be derived one relatively to $\{N'^i\}$. All that is necessary is to construct $N'^{j_1} \supset N^{i_1} \supset N'^{j_2} \supset N^{i_2} \cdots$ and take $\Gamma'^h = \Gamma^{i_r}$, $j_{r-1} < h \leqq j_r$; $j_0 = 1$.

As a matter of notation we shall always designate a cycle on G by replacing in the generic term Γ^i by γ. Thus $\{\Gamma_\mu^{ij}\}$ will be designated by γ_μ^j, etc.

25. **Homologies and related invariants.** We write down by definition $\gamma_\mu \approx 0 \bmod. G$ whenever for every p there is a q such that for $i \geq q$, $\Gamma^i \approx 0 \bmod. N^p$. It is a simple matter to verify that if this condition is fulfilled by the given representation $\{\Gamma^i\}$ relatively to $\{N^i\}$ it is likewise for any subsequence $\{\Gamma^{s_i}\}$ as a representation relatively to $\{N^{s_i}\}$ or in fact as a representation relatively to any other $\{N'^i\}$ to which it

may belong. The converse follows also at once by making use of the definition of cycles.

Let γ_μ^j, $j = 1, 2, \cdots, r$, be several cycles on G, with representations $\{\Gamma_\mu^{ij}\}$ relatively to $\{N^i\}$. Then $\{\sum \lambda_j \Gamma_\mu^{ij}\}$ is also the representation of a cycle relatively to $\{N^i\}$. If $\{N^i\}$ is replaced by $\{N'^i\}$ the representation of the same cycle preserves its form: Namely if γ^j is represented by $\{\Gamma_\mu'^{ij}\}$ then the new cycle is represented by $\{\sum \lambda_j \Gamma_\mu'^{ij}\}$. We are therefore justified in denoting it by $\sum \lambda_j \gamma_\mu^j$. When there exists a non trivial sum of this nature ≈ 0 mod. G the γ's are *dependent* mod. G, otherwise they are independent. The maximum number of independent μ-cycles of G is as usual denoted by $R_\mu(G)$.

Let Γ_μ be a cycle of M_n. We shall define $\gamma_\mu \approx \Gamma_\mu$ mod. M_n, as meaning that for every i above a certain limit $\Gamma_\mu^i \approx \Gamma_\mu$ mod. M_n. In particular $\gamma_\mu \approx 0$ mod. M_n whenever under the same conditions $\Gamma_\mu^i \approx 0$ mod. M_n. These definitions are again manifestly independent of the representation. They make it possible to consider mixed homologies involving cycles of both M_n and G. There follows from them a precise meaning for the numbers $r_\mu(G)$, $S_\mu(M_n - G)$, $T_\mu(G)$. We have the following noteworthy properties.

I. *r, S, T are simultaneous topological invariants of M_n and G. R is a topological invariant of G alone.*

The property of R is the only one which is not an immediate consequence of the definitions. To prove it we observe that by reasoning as in No. 12 we can show that for i sufficiently great any $\Gamma_\mu \approx$ mod. some N^{i-h} to a Γ_μ' with its vertices on G. More generally if N^i carries a $C_{\mu+1} \to \Gamma_\mu$, $C_{\mu+1}$ is turned into a similar $C'_{\mu+1}$ on an N^{i-h} and with its vertices on G, whose boundary is the Γ' associated with Γ. In both cases $i - h$ tends to ∞ with i. By taking advantage of this we can limit ourselves to γ_μ's whose sequences $\{\Gamma_\mu^i\}$ consist of cycles with vertices on G, with the added restriction that when $\gamma_\mu \approx 0$, the complexes bounded by the terms for i above a certain limit also have their vertices on G. But if $G \subset M_n$ is homeomorphically transformed into $G' \subset M_n'$, we may establish a one-one correspondence between such specialized sequences associated with suitable neighborhood sequences on M_n and M_n', and of such nature that sums and multiples of cycles go into sums and multiples of the corresponding cycles, and cycles ≈ 0 mod. G go into cycles ≈ 0 mod. G' and conversely. This is sufficient to prove that R_μ is a topological invariant of G alone.

The above reasoning is essentially the same as used by Brouwer and Vietoris in a similar order of ideas. For further details see their papers quoted in the Introduction (Nos. 10,1).

II. *When $G = C_k$ the four invariants become the invariants defined in the usual manner, i. e. with the "on" relation as ordinarily understood.* For given γ_μ, when i exceeds a certain limit $\Gamma^i \approx \bar{\Gamma}^i$, mod. N^{i-h}, $\bar{\Gamma}^i \subset C_k$; given p sufficiently great there is a q such that for $i, j \geqq q$ $\bar{\Gamma}^i \approx \bar{\Gamma}^j$ mod. N^p, hence mod. C_k (reduction theorem). It follows that γ_μ identifies a unique Γ_μ on C_k; conversely to Γ_μ on C_k corresponds γ_μ with every $\Gamma^i = \Gamma$. That the numbers in question are unchanged under the correspondence plus the change in the *on* relation is shown for example as in the similar case of I.

26. **Relation between the cycles very near G and the cycles on G. Reduction theorem.** We have repeatedly pointed out the importance played in the treatment of C_k by the reduction theorem of No. 12 for the cycles very near C_k. We shall now show that if $\{N^i\}$ is properly chosen a similar result holds for G.

Consider first three neighborhoods $N \supset N' \supset N''$. A set of μ-cycles on N' independent mod. N is likewise on N' independent mod. N'. Hence the maximum number ϱ of μ-cycles of N' independent mod. $N \geqq$ the same number for $N'' \subset N'$. Let R^N be its minimum: If $\varrho = R^N_\mu$ already for N', it will assume the same value for any neighborhood $\subset N'$.

Since a set of cycles of N'' independent mod. N, is necessarily independent mod. N' as well, $R^N \leqq R^{N'}$. Hence R^N does not decrease when we pass from N to any $N' \subset N$. We shall denote its maximum by R'_μ and return to it in a moment.

Take now $\{N^i\}$ such that $N^{i+1} \supset$ precisely R^{N^i} cycles independent mod. N^i. As observed above, since $N^{i+2} \subset N^i$ it will carry the same maximum number R^{N^i} of cycles independent mod. N^i as N^{i+1}. Hence every Γ_μ of $N^{i+1} \approx$ mod. N^i to a Γ^{i+2}_μ on N^{i+2}. The same reasoning may be continued and we shall thus have a sequence $\gamma_\mu = \{\Gamma^h_\mu\}$; $\Gamma^h = \Gamma^{i+1}$, $h \leqq i+1$; with the property $\Gamma^h \approx \Gamma^{h+1}$, mod. N^{h-1}. It defines then a unique cycle on G and by its very construction $\gamma_\mu \approx \Gamma_\mu$, mod. M_n. But Γ_μ is after all any cycle of M_n whose points are sufficiently near G. Hence

Reduction Theorem: *Any Γ_μ of M_n on a neighborhood of G whose width is sufficiently small, is $\approx$ a cycle on G mod. M_n.*

Remark. We can extract from any sequence $\{N^i\}$ a subsequence $\{N^{s_i}\}$ having the special property of the sequence considered above. Therefore we can always construct a sequence $\{N^i\}$ having that property and with every N^i polyhedral.

27. *In the preceding proof fractionary cycles played an essential part.* Suppose indeed that Γ_μ were an integral cycle which is not an actual restriction in view of the nature of the theorem sought. Then we can only assert in general that it is $\approx$ a fractionary cycle Γ^{i+2}. Or remaining in the domain of integral cycles we can only assert that there

is an integer $q_{i+1} \neq 0$ such that $q_{i+1} \Gamma_\mu \approx \Gamma_\mu^{i+2}$, mod. N^i. Similarly $q_{i+1} \Gamma_\mu^{i+2} \approx \Gamma_\mu^{i+3}$, mod. N^{i+1}, the new cycle being again integral, etc. If the sequence of q's is bounded we can find a $\{\Gamma_\mu^i\} \approx$ a multiple of Γ_μ, and with terms all integral cycles, but not otherwise, and the theorem of reduction would fail to hold. These sequences of integral cycles give rise to the cycles introduced by Vietoris.

Another way of looking at the matter is this: A Γ_μ may be such that a multiple $q\Gamma_\mu \approx$ a cycle on a neighborhood of width ε of G; if q remains bounded as ε tends to zero Vietoris-cycles (sequences of integral cycles) are adequate. If not our more general type must be used.

28. In his paper (footnote 18) Vietoris indicates that Brouwer introduced a number which is R'_μ of No. 25 (maximum of R_μ^N), his procedure being substantially the same as ours, i. e, without making use of any entities definable as cycles on G.* Let us show that its value is $R_\mu(G)$. Let indeed N^{i+1} carry r cycles independent mod. N^i: $\Gamma_\mu^{h+1,j}$, $j = 1, 2, \cdots, r$. By the above process they give rise to r cycles on G, γ_μ^j, certainly independent mod. G. For if $\sum \lambda_j \gamma^j \approx 0$ mod. G there is a certain p which we may take $> h$, such that for $i > p$, $\sum \lambda_j \Gamma^{ij} \approx 0$, mod. N^p, hence mod. N^h, since $N^p \subset N^h$. For a similar reason $\Gamma^{ij} \approx \Gamma^{h+1,j}$, mod. N^h. Hence $\sum \lambda_j \Gamma^{h+1,j} \approx 0$, mod. N^h, contrary to assumption. Therefore the γ's are independent and $R_\mu(G) \geqq r$, hence $\geqq R'_\mu$ maximum of r.

On the other hand if we possess r independent cycles γ_μ^j on G, the r cycles $\Gamma_\mu^{i+1,j}$ are on N^i and independent mod. a certain $N^p \subset N^i$, hence mod. N^i also. Therefore $r \leqq R'_\mu$ maximum of R^{N_i} and consequently $R_\mu(G) \leqq R'_\mu$. This proves that $R'_\mu = R_\mu(G)$ as asserted.

29. **Cycles relatively to G.** By definition a μ-cycle rel. G or a complex with boundary on G, is a sequence of complexes $\delta_\mu = \{\Delta_\mu^i\}$, with $\Delta^i \to \Gamma_{\mu-1}^i \subset N^i$ and furthermore such that for every p there is a q having the property that for $i, j \geqq q$, $\Delta^i - \Delta^j \subset N^p$. From this follows $\Gamma^i - \Gamma^j \approx 0$ mod. N^p, so that $\gamma_{\mu-1} = \{\Gamma_{\mu-1}^i\}$ is a cycle on G, obviously ≈ 0 mod. M_n. We express the relation between γ and δ by $\delta_\mu \to \gamma_{\mu-1}$.

In regard to notation we make the same convention as for cycles except that δ will replace γ.

As in the case of cycles we agree that when two sequences $\{\Delta_\mu^i\}$, $\{\Delta_\mu'^i\}$ with the above properties relatively to two sequences $\{N^i\}$, $\{N'^i\}$, have

* I was not aware of this until the present paper was finished. The relation $R'_\mu = R_\mu(G)$ to be proved presently may then be stated thus: Brouwer's generalized Betti numbers or connection indices are the same as those that we have obtained by means of our sequences of fractionary cycles. It may be observed that, when dealing with cycles mod. 2, Vietoris-cycles are sufficient and do not differ from our type.

an infinite subsequence of common elements, they define one and the same δ_μ of which they are distinct representations. Then the corresponding sequences $\{\Gamma^i_{\mu-1}\}$, $\{\Gamma'^i_{\mu-1}\}$ have also an infinite subsequence of common elements and therefore are representations of the same $\gamma_{\mu-1}$. Hence the boundary relation $\delta_\mu \to \gamma_{\mu-1}$ is independent of the representations of γ or δ.

As already observed from $\delta_\mu \to \gamma_{\mu-1}$ follows $\gamma_{\mu-1} \approx 0$ mod. M_n. Conversely to any $\gamma_{\mu-1} \approx 0$ mod. M_n corresponds a relative cycle $\delta_\mu \to \gamma_{\mu-1}$. For since $\gamma \approx 0$ mod. M_n there is an integer α such that for every $i \geqq \alpha$: (a) there is a p_i tending to ∞ with i such that $\Gamma^i \approx \Gamma^{i+1}$ mod. N^{p_i}. We specifically denote by p_i the greatest integer having that property. (b) $\Gamma^i \approx 0$ mod. M_n. In view of (a) N^{p_i} carries a complex $C^i_\mu \to \Gamma^{i+1}_{\mu-1} - \Gamma^i_{\mu-1}$. Also owing to (b), M_n carries a relative cycle $\Delta^\alpha_\mu \to \Gamma^\alpha_{\mu-1}$. Let now

$$\Delta^i_\mu = \Delta^\alpha_\mu + \sum_{h=\alpha}^{i-1} C^i_\mu; \qquad i > \alpha.$$

This defines a δ_μ since it assigns the Δ's for an infinite sequence of N's. Since $\Delta^i_\mu \to \Gamma^i_{\mu-1}$ $i > \alpha$, $\delta_\mu \to \gamma_{\mu-1}$ and it is therefore the required relative cycle.

30. The following definitions will enable us to consider relative boundary relations and homologies:

(a) We define $\delta_\mu \approx 0$, mod. M_n rel. G as meaning that for every p there is a q such that when $i \geqq q$, $\Delta^i_\mu \approx 0$ mod. M_n rel. N^p.

(b) Let $\delta^j_\mu \to \gamma^j_{\mu-1}$, $j = 1, 2, \cdots, \nu$. Then $\{\sum \lambda_j \Delta^{ij}_\mu\}$ is a δ_μ to be denoted by $\sum \lambda_j \delta^j_\mu$ and we verify at once that $\sum \lambda_j \delta^j_\mu \to \sum \lambda_j \gamma^j_{\mu-1}$. The δ's are dependent mod. M_n rel. G whenever a non trivial sum

$$\sum \lambda_j \delta^j_\mu \approx 0 \text{ mod. } M_n \text{ rel. } G.$$

With p, q, as above it means that for $i > q$,

$$\sum \lambda_j \Delta^{ij} \approx 0, \text{ mod. } M_n \text{ rel. } N^p.$$

When no such condition is fulfilled the relative cycles δ^j_μ are independent mod. M_n or mod. M_n rel. G respectively. The maximum number of δ's independent mod. M_n rel. G is to be denoted according to our conventions by $R_\mu(M_n; G)$.

(c) In regard to relations between the δ's and ordinary cycles of M_n, we identify any Γ_μ of M_n, with the relative cycle $\{\Delta^i_\mu\}$ such that every $\Delta^i = \Gamma$. Thus the homology

$$\sum \lambda_j \Gamma^j_\mu + \sum \nu_j \delta^j_\mu \approx 0, \text{ mod. } M_n \text{ rel. } G,$$

means that with p, q as above, for $i \geqq q$,

$$\sum \lambda_j \Gamma^j + \sum \nu_j \Delta^{ij} \approx 0, \text{ mod. } M_n, \text{ rel. } N^p.$$

The preceding definitions are freed from the particular $\{N^i\}$ chosen in precisely the same manner as in the analogous cases of the cycles. It follows in particular that $R_\mu(M_n; G)$ is a simultaneous topological invariant of M_n and G.

31. **Kronecker indices and intersection theorems.** We denote henceforth by K^i, Φ^i, the K and Φ of No. 24 corresponding to N^i. Let $\Gamma_\mu \subset M_n - G$. There exists then a p such that $\Gamma_\mu \subset K^p - \Phi^p = M_n - N^p$. To p there corresponds a q such that $\Delta^i - \Delta^j \subset N^p$ for $i, j \geqq q$. Then $(\Gamma \cdot \Delta^i - \Delta^j) = 0$ since the two complexes have no common point. Therefore $(\Gamma \cdot \Delta^i) = (\Gamma \cdot \Delta^j)$. Thus $(\Gamma \cdot \Delta^i)$ has a value independent of i provided i exceeds a certain limit, and this value is that we define as the Kronecker-index $(\Gamma_\mu \cdot \delta_{n-\mu})$. If we were to take in place of the Δ's any infinite subsequence of them we would arrive at the same index. From this we conclude immediately that it is independent of the special $\{N^i\}$ and the corresponding representation of δ.

Having shown by this time that all notions introduced are independent of $\{N^i\}$, it will be convenient to take for it a sequence with the special property of No. 26 and with every N^i polyhedral. This is always possible and will imply no loss of generality. See in this connection the Remark at the end of No. 26.

Having defined the index $(\Gamma \cdot \delta)$, we can state the same intersection theorems as for C_k and we proceed to prove them.

I. *In order that* $\Gamma_\mu \approx 0$ *mod.* $M_n - G$ *it is necessary and sufficient that* $(\Gamma_\mu \cdot \delta_{n-\mu}) = 0$ *for every* $\delta_{n-\mu}$.

(a) Let $\Gamma_\mu \approx 0$ mod. $M_n - G$, so that $M_n - G$ carries a $C_{\mu+1} \to \Gamma_\mu$. There exists an α such that $C_{\mu+1} \subset K^{\alpha-1} - \Phi^{\alpha-1}$. Then the boundary of $\Delta^{\alpha+i}$, which is on N^α, does not meet C. Hence (No. 14), $(\Gamma \cdot \Delta^{\alpha+i}) = 0$ for every i, which means by definition $(\Gamma \cdot \delta) = 0$. The condition is thus necessary.

(b) Let now, Γ_μ being fixed, $(\Gamma_\mu \cdot \delta) = 0$ for every δ. There exists an α such that Γ_μ is wholly exterior to N^α. Let $\Delta_{n-\mu}$ be an arbitrary cycle rel. N^α. It is sufficient to show that from $(\Gamma_\mu \cdot \delta) = 0$, follows $(\Gamma_\mu \cdot \Delta) = 0$, for as a consequence of the theorem as proved for C_k we have then $\Gamma_\mu \approx 0$ mod. $M_n - N^\alpha$, hence also mod. $M_n - G$.

To prove that $(\Gamma_\mu \cdot \Delta) = 0$, we observe that if $\Delta \to \Gamma_{n-\mu-1} \subset N^\alpha$ in view of the reduction theorem of No. 26 we are assured of the existence of a $\gamma_{n-\mu-1}$ such that $\Gamma^i_{n-\mu-1} = \Gamma_{n-\mu-1}$, $i \leqq \alpha$, and that $\Gamma^{i+1}_{n-\mu-1} \approx \Gamma^i_{n-u-1}$

mod. N^{i-1} whatever i. Since $\Gamma_{n-\mu-1} \approx 0$ mod. M_n, likewise $\gamma_{n-\mu-1} \approx 0$ mod. M_n. Hence there exists a $\delta_{n-\mu} \to \gamma$, and as shown in No. 29, δ can be so chosen that for $i > \alpha$ $\Delta^i = \Delta + C_{n-\mu}$, $C_{n-\mu} \subset N^\alpha$, for the integer p_i of No. 29 is here $i-1$. But since Γ_μ has no point on N^α it will not meet C, hence $(\Gamma_\mu \cdot C) = 0$. But the assumption $(\Gamma_\mu \cdot \delta) = 0$ is equivalent to $(\Gamma_\mu \cdot \Delta^i) = 0$ for i sufficiently great. It follows $(\Gamma_\mu \cdot \Delta) = (\Gamma_\mu \cdot \Delta_i - C)$ $= (\Gamma_\mu \cdot \Delta^i) - (\Gamma_\mu \cdot C) = 0$, and this as we have observed completes the proof of I.

II. *In order that* $\delta_{n-\mu} \approx 0$ mod. M_n rel. G *it is necessary and sufficient that* $(\Gamma_\mu \cdot \delta) = 0$ *for every* $\Gamma_\mu \subset M_n - G$.

(a) Let $\delta_{n-\mu} \approx 0$ mod. M_n rel. G. Given any particular Γ_μ let α again be such that Γ_μ has no points on N^α. For i above a certain limit, $\Delta^i \approx 0$ mod. M_n rel. N^α. Hence by the theorem for C_k (here N^α), $(\Gamma_\mu \cdot \Delta^i) = 0$, which means $(\Gamma_\mu \cdot \delta) = 0$. The condition is therefore necessary.

(b) Let $(\Gamma \cdot \delta) = 0$ whatever Γ. Let p be arbitrary and Γ fixed on K^p. There exists a $q > p$ such that $(\Gamma \cdot \Delta^i) = 0$ for $i \geqq q$. Let Γ_μ^j, $j = 1, 2, \cdots$, $R_\mu(K^p)$ be a maximal set of independent cycles on K^p. Then $\Gamma \approx \sum \lambda_j \Gamma^j$, mod. K^p. For each Γ^j there is a q_j analogous to the above q for Γ. Let q' be their maximum. Then for $i \geqq q'$, every $(\Gamma^j \cdot \Delta^i) = 0$, hence $(\Gamma \cdot \Delta^i) = \sum \lambda_j (\Gamma^j \cdot \Delta^i) = 0$. Hence by II as proved for an M_n and C_k with $C_k = N^p$, $\Delta^i \approx 0$, mod. M_n rel. N^p, which means $\delta \approx 0$ mod. M_n rel. G and completes the proof of II.

32. **Proof of formulas (4), (5), (6).** The proofs of (4) and (5) are exactly as for C_k, (Nos. 20, 21) except that N, K, Φ corresponding to ε sufficiently small, are to be replaced by N^i, K^i, Φ^i, i sufficiently great. There is no need to repeat them here. Slight modifications are necessary in the derivation of (6). We have as in No. 22 three sets independent mod. M_n rel. G: Γ_μ^j, D_μ^j, $\delta_\mu^j \to \gamma_{\mu-1}^j$ with the γ's a maximal set of $\mu - 1$ cycles independent mod. G but ≈ 0 mod. M_n and we must show their independence.

Assume

$$\sum \lambda_j \Gamma^j + \sum \nu_j D^j + \sum \varrho_j \delta^j \approx 0, \text{ mod. } M_n \text{ rel. } G.$$

This means that for every p there is a q such that for $i \geqq q$

$$\sum \lambda_j \Gamma^j + \sum \nu_j D^j + \sum \varrho_j \Delta^{ij} \approx 0, \text{ mod. } M_n \text{ rel. } N^p.$$

There exists an α such that every $\Gamma^j \subset K^\alpha - \Phi^\alpha$. Choose then p such that $N^p \subset N^\alpha - \Phi^\alpha$. By taking the intersections with Φ^α we derive from the last homology

$$\sum \nu_j D^j \cdot \Phi^\alpha + \sum \varrho_j \Delta^{ij} \cdot \Phi^\alpha \approx 0, \text{ mod. } \Phi^\alpha.$$

Observing as in No. 22 that $D^j \cdot \Phi^\alpha \approx \Delta^{ij} \cdot \Phi^\alpha - \Gamma^{ij} \approx 0$ mod. N^α, since they bound respectively the parts of D^j and Δ^{ij} on N^α, we find $\sum \varrho_j \Gamma^{ij} \approx 0$, mod. N^α, $i \geqq \alpha + 1$. This means by definition $\sum \varrho_j \gamma^j \approx 0$, mod. G contrary to assumption. Therefore every $\varrho = 0$. The rest of the proof of the independence of the three sets is as for C_k.

The second part of the proof requires that we show that every relative cycle δ_μ depends upon those of the given three sets. In view of the maximal property of the cycles γ_μ^j, we have $\gamma \approx \sum \varrho_j \gamma^j$ mod. G. For every p there exists then a q such that when $i \geqq q$, $\Gamma_{\mu-1}^i - \sum \varrho_j \Gamma_{\mu-1}^{ij} \approx 0$, mod. N^p. Therefore N^p carries a $C_\mu \to \Gamma^i - \sum \varrho_j \Gamma^{ij}$. Hence $\Delta^i - \sum \varrho_j \Delta^{ij} - C \to 0$. The cycle of M_n so obtained depends mod. M_n rel. N^p upon the cycles of the two sets Γ^j, D^j. Therefore for every $i > q$

$$\Delta^i \approx \sum \lambda_j \Gamma^j + \sum \nu_j D^j + \sum \varrho_j \Delta^{ij}, \text{ mod. } M_n \text{ rel. } N^p$$

which in effect means that

$$\delta \approx \sum \lambda_j \Gamma^j + \sum \nu_j D^j + \sum \varrho_j \delta^j, \text{ mod. } M_n \text{ rel. } G$$

and as in No. 22 completes the proof of (6).

33. (Added in proof.) I have recently obtained a noteworthy extension of (5) which has the advantage of being more symmetrical and of requiring a single intersection theorem. Given F, G, closed subsets of M_n, the formula reads

$$R_\mu (M_n - F; G) = R_{n-\mu} (M_n - G; F).$$

The proof entirely analogous to that of (4) or (5) need not be given here.

I am indebted to Alexandroff for the following references to papers in the Göttinger Nachrichten: (a) Alexandroff, Nov. 25, (1927). (b) Pontrjagin, same date. (c) Pontrjagin presented in March 1928. We may also mention a paper (d) Frankl, Wiener Berichte, December 1927, pp. 689–699. All these papers deal with duality relations but only for invariants mod. 2. The argument in (b), (c), (d) is substantially the same as ours. In Pontrjagin's second paper alone is the containing M_n not an S_n. See also a closely related paper by Alexandroff, Fundamenta Mat., vol. 11 (1928), pp. 222–228.

ON RAYLEIGH'S PRINCIPLE IN THE THEORY OF DIFFERENTIAL EQUATIONS OF MATHEMATICAL PHYSICS AND ON EULER'S METHOD IN CALCULUS OF VARIATIONS.*

By N. Bogoliouboff and N. Kryloff.†

It is well known from mechanical theories that the passage from discrete systems to continuous ones can be made in two different ways. Thus in the early days of the Calculus the mathematicians of the XVIII century carried out the limiting process in the equations of motion of a discrete system and obtained in this way the differential equations of the Mathematical Physics, which can be considered as the equations of motion of the limiting continuous system. On the other hand the limiting process can be carried out in the solutions of the corresponding equations in finite differences. The question arises now if these two different limiting processes will lead to the same result.

Guided by their sense of intuition the mathematicians and above all the physicists have answered this delicate question in the affirmative. Such eminent scientists as Lord Rayleigh, H. Poincaré and others have used this principle, the so called Rayleigh's Principle, as a heuristic method in several important problems of Mathematical Physics.

The field of the mathematical investigation hereby opened, is far from being completely explored, as the reader may easily convince himself by consulting e. g. the article by M. Bôcher on the work of Sturm‡ and also the recent paper by R. D. Carmichael.§ A recent paper by M. Plancherel‖ is devoted to the justification of Rayleigh's Principle, but the reasoning of this author as well as that of some other writers who have treated of the same question, as far as we can see, does not enable to estimate the

* Received July 2, 1926.

† A discussion which I had with the eminent English scientist Prof. L. Richardson, on board of the "Caronia" on my way to the International Mathematical Congress at Toronto, 1924, led my attention to this problem. The solution is given in the present paper written in collaboration with my pupil Mr. N. Bogoliouboff. N. Kryloff.

‡ The published and unpublished work of Ch. Sturm on algebraic and differential equations, Bull. Amer. Math. Soc., vol. 18 (1912), pp. 1–18.

§ Boundary value problems, Amer. Jour. Math., vol. 44 (1922), pp. 129–152.

‖ Le passage à la limite des équations aux différences aux équations différentielles dans les problèmes aux limites, Bull. des sci. math., (2), vol. 47 (1923), pp. 155–160, 170–177.

order of the error committed at the n-th step of approximation. The same observation may be made concerning a paper by R. B. Robbins* which is devoted to a detailed study of the equations of the Calculus of Variations from Euler's point of view.

The estimate of the error is not without some interest, however, especially from the point of view of practice, because in several problems of engineering one has to replace the differential equation of the problem by an equation in finite differences and to solve this last one by algebraic methods. For this reason we shall present in this paper some methods for estimating the error in question, restricting ourselves to the simplest cases, as generalizations are easily made.

The authors express their sincere thanks to Prof. Dr. Tamarkin, of the Brown University, for his assistance during the revision of this paper.

1. A special differential equation of the second order. We shall consider the differential system:

$$\frac{d^2y}{dx^2} = q(x)\,y(x)+f(x), \qquad q(x)>0, \tag{1}$$
$$y(0) = y(1) = 0$$

and we are going to be concerned with the order of the difference

$$\delta_{k,n} = y(x_k) - y_k^{(n)} = y_k - y_k^{(n)}, \qquad x_k = \frac{k}{n} \qquad (k = 0, 1, \cdots, n),$$

where $y_k^{(n)}$ is the solution of the corresponding system in finite differences:

$$\frac{\Delta^2 y_{k-1}^{(n)}}{\Delta x^2} = q_k\,y_k^{(n)}+f_k; \quad q_k = q(x_k), \quad f_k = f(x_k); \quad \Delta x = \Delta = \frac{1}{n}; \tag{2}$$
$$y_0^{(n)} = y_n^{(n)} = 0 \qquad (k = 1, 2, \cdots, n-1).$$

It is readily seen that the Green's function

$$G(x, t) = \begin{cases} (1-x)\,t, & \text{if} \quad t \leqq x, \\ (1-t)\,x, & \text{if} \quad t \geqq x \end{cases}$$

of the system

$$\frac{d^2y}{dx^2} = 0, \qquad y(0) = y(1) = 0, \tag{1_1}$$

coincides, at the points x_i, $t_k = k/n$, with the Green's function†

* A method in the calculus of variations, Amer. Jour. Math., vol. 37 (1915), pp. 307–344.

† Plancherel, loc. cit., pp. 158–160.

$$G_{ik}^{(n)} = G_{ik} = \begin{cases} \left(1 - \frac{i}{n}\right)\frac{k}{n}, & \text{if } k \leqq i, \\ \left(1 - \frac{k}{n}\right)\frac{i}{n}, & \text{if } k \geqq i \end{cases}$$

of the system

$$(2_2) \qquad \frac{\Delta^2 y_{k-1}^{(n)}}{\Delta x^2} = 0, \qquad y_0^{(n)} = y_n^{(n)} = 0.$$

But systems (1) and (2) are equivalent respectively to the equations

$$(3) \qquad y(x) = -\int_0^1 G(x, t)\, q(t)\, y(t)\, dt + F(x),$$

$$(4) \qquad y_k^{(n)} = -\sum_{i=0}^{n} G_{ki}\, q_i\, y_i^{(n)}\, \Delta + F_k^{(n)}$$

where

$$F(x) = -\int_0^1 G(x, t)\, f(t)\, dt,$$

$$F_k^{(n)} = -\sum_{i=0}^{n} G_{ki}\, f_i\, \Delta.$$

Hence

$$(5) \qquad \delta_{k,n} = -\sum_{i=0}^{n} G_{ki}\, q_i\, \delta_{i,n}\, \Delta + \varepsilon_{k,n}$$

where

$$\varepsilon_{k,n} = -\left[\sum_{i=0}^{n} G_{ki}\, q_i\, y_i\, \Delta - \int_0^1 G(x, t)\, y(t)\, dt\right] + [F_k^{(n)} - F_k] = O\left(\frac{1}{n^2}\right),$$

if we assume that $q(x)$ and $f(x)$ possess first derivatives which satisfy a Lipschitz's condition on (0, 1). This follows immediately from the fact that the sums

$$\sum_{i=0}^{n} G_{ki}\, q_i\, y_i\, \Delta, \qquad \sum_{i=0}^{n} G_{ki}\, f_i\, \Delta,$$

by virtue of the boundary conditions, reduce to the trapezoidal formulas for the integrals

$$\int_0^1 G(x, t)\, q(t)\, y(t)\, dt, \qquad \int_0^1 G(x, t)\, f(t)\, dt$$

respectively, and the remainder of the trapezoidal formula is of the order $O\left(\frac{1}{n^2}\right)$, under our conditions.

On multiplying (5) by $q_k\, \delta_{k,n}\, \Delta$ and summing over k we obtain

$$(6) \qquad \sum_{k=0}^{n} q_k\, \delta_{k,n}^2\, \Delta = -\sum_{i,k=0}^{n} G_{k,i}\, q_i\, q_k\, \delta_{i,n}\, \delta_{k,n}\, \Delta^2 + \sum_{k=0}^{n} \varepsilon_{k,n}\, q_k\, \delta_{k,n}\, \Delta.$$

By a method analogous to that used in the theory of the integral equations in demonstrating the definitely positive character of the kernel $G(x, t)$, we can prove that

$$\sum_{k,i=0}^{n} G_{ki}\, h_k\, h_i \geqq 0$$

for an arbitrary sequence $h_0, h_1, \cdots, h_n$. From (6) it is readily seen then:

$$\sum_{k=0}^{n} q_k\, \delta_{k,n}^2\, \Delta \leqq \sum_{k=0}^{n} \varepsilon_{k,n}\, q_k\, \delta_{k,n}\, \Delta,$$

whence, by an easy application of Cauchy's inequality:

$$\sum_{k=0}^{n} q_k\, \delta_{k,n}^2\, \Delta \leq \max \varepsilon_{k,n}^2 \cdot \sum_{k=0}^{n} q_k\, \Delta.$$

Substituting in (5) and applying Cauchy's inequality again, we find:

$$\delta_{k,n} = y_k^{(n)} - y_k = O(\max |\varepsilon_{k,n}|) = O\left(\frac{1}{n^2}\right).$$

Thus we obtain an estimate of the order of the error $\delta_{k,n}$ and at the same time a new proof of Rayleigh's Principle.

2. **Examination of the general case.** We consider now the differential system

$$\begin{gathered} \frac{d^2 y(x)}{d x^2} + \lambda\, q(x)\, y(x) = f(x), \\ y(0) = y(1) = 0, \end{gathered} \tag{7}$$

together with the corresponding system in finite differences:

$$\begin{gathered} \frac{\Delta^2 y_{k-1}^{(n)}}{\Delta x^2} + \lambda\, q_k\, y_k = f_k \qquad (k = 1, 2, \cdots, n-1), \\ y_0^{(n)} = y_n^{(n)} = 0. \end{gathered} \tag{8}$$

In the same manner as before we obtain

$$\delta_{k,n} = y_k^{(n)} - y_k = \lambda \sum_{i=0}^{n} G_{k,i}\, q_i\, \delta_{i,n}\, \Delta + \varepsilon_{k,n}$$

where

$$\varepsilon_{k,n} = \lambda \left[\sum_{i=0}^{n} G_{ki}\, q_i\, y_i\, \Delta - \int_0^1 G(x_k, t)\, q(t)\, y(t)\, dt \right] + [F_k^{(n)} - F^k].$$

A reasoning analogous to that which yields Schmidt's well known formula for the solution of a non-homogeneous integral equation, enables us to write $\delta_{k,n}$ in the form:

$$\delta_{k,n} = \varepsilon_{k,n} - \lambda \sum_{\nu=1}^{n-1} \frac{\varepsilon_\nu \varphi_{\nu k}^{(n)}}{\lambda - \lambda_{\nu,n}}; \quad \varepsilon_\nu = \sum_{i=0}^{n} q_i \varepsilon_{i,n} \varphi_{\nu i}^{(n)} \Delta, \tag{9}$$

where $\varphi_{\nu,k}^{(n)}$ are fundamental solution of system (8) that is to say

$$\begin{aligned} &\frac{\Delta^2 \varphi_{\nu k-1}^{(n)}}{\Delta x^2} + \lambda_{\nu,n} q_k \varphi_{\nu,k}^{(n)} = 0 \\ &\varphi_{\nu,0}^{(n)} = \varphi_{\nu,n}^{(n)} = 0 \end{aligned} \qquad (\nu = 1, 2, \cdots, n-1; \ k = 0, 1, \cdots, n), \tag{8_1}$$

$\lambda_{\nu,n}$ being the characteristic values of the problem (8) in finite differences, or simply "difference characteristic values". It is convenient to note that

$$\sum_{i=0}^{n} a_{\nu,i}^{(n)} \varphi_{\mu,i}^{(n)} q_i \Delta = \begin{cases} 0, & \text{if } \mu \neq \nu, \\ 1, & \text{if } \mu = \nu, \end{cases} \tag{10}$$

$$G_{ik} = \sum_{\nu=1}^{n-1} \frac{\varphi_{\nu,i}^{(n)} \varphi_{\nu,k}^{(n)}}{\lambda_{\nu,n}}, \tag{11}$$

$$\begin{aligned} &f_k = \sum_{\nu=1}^{n-1} f^{(\nu)} \varphi_{\nu,k}^{(n)} \quad \text{where} \quad f(\nu) = \sum_{k=0}^{n} f_i \varphi_{\nu,i}^{(n)} q_i \Delta; \\ &\sum_{\nu=1}^{n-1} f^{(\nu)2} \leqq \sum_{k=0}^{n} f_k^2 q_k \Delta, \end{aligned} \tag{12}$$

which may be proved in exactly the same manner as the corresponding formulas in the theory of the integral equations.

Then from (9) it is readily seen:

$$\begin{aligned} \sum_{k=0}^{n} q_k \delta_{k,n}^2 \Delta &\leqq 2 \sum_{k=0}^{n} q_k \varepsilon_{k,n}^2 \Delta + 2\lambda^2 \sum_{k=0}^{n} q_k \Delta \left(\sum_{\nu=1}^{n-1} \frac{\varphi_{\nu,k}^{(n)} \varepsilon_\nu}{\lambda - \lambda_\nu} \right)^2 \\ &\leqq 2 \sum_{k=0}^{n} q_k \varepsilon_{k,n}^2 \Delta + \frac{2\lambda^2}{d_n^2} \sum_{\nu=1}^{n-1} \varepsilon_\nu^2 \leqq \frac{M}{d_n^2} \sum_{k=0}^{n} q_k \varepsilon_{\nu,k}^2 \Delta = \frac{1}{d_n^2} O(\max \varepsilon_{k,n}^2), \end{aligned} \tag{13}$$

where d_n denotes $\min |\lambda - \lambda_{\nu,n}|$, and M is a numerical factor which does not depend on n. We shall prove below that the difference between two corresponding characteristic values λ_ν, $\lambda_{\nu,n}$ of the systems (7) and (8) is of the order $O\left(\frac{1}{n}\right)$. Hence, observing that $\lambda - \lambda_{\nu,n} = (\lambda - \lambda_\nu) + (\lambda_\nu - \lambda_{\nu,n})$, from (13) we see at once that

$$\sum_{k=0}^{n} q_k \delta_{k,n}^2 \Delta = \frac{1}{d^2} O(\max \varepsilon_{k,n}^2),$$

where d denotes $\min |\lambda - \lambda_\nu|$. We thus obtain the desired estimate of the error $\delta_{k,n}$ because the same reasoning as that used at the end of § 1 shows:

$$\delta_{k,n} = \frac{1}{d} O(\max |\varepsilon_{k,n}|) = \frac{1}{d} O\left(\frac{1}{n^2}\right).$$

3. Calculation of characteristic values and of fundamental functions. It is easy to prove, exactly as in the theory of the integral equations, that the first difference characteristic value is the reciprocal of the maximum of the double sum

$$\sum_{i,j=0}^{n} K_{ij}\, h_i\, h_j\, \Delta^2, \quad K_{ij} = G_{ij}\, q_i\, q_j, \tag{14}$$

where $h_0, h_1, \cdots, h_n$ is an arbitrary sequence satisfying the condition

$$\sum_{i=0}^{n} q_i\, h_i^2\, \Delta = 1. \tag{15}$$

The sequence $h_0^{(n)}, \cdots, h_n^{(n)}$ which maximizes (14) is uniquely determined and coincides with the normalized fundamental solution of the difference boundary problem (8) that is to say:

$$\frac{1}{\lambda_{1,n}} = \max \sum_{i,j=0}^{n} K_{ij}\, h_i\, h_j\, \Delta^2 = \sum_{i,j=0}^{n} K_{ij}\, h_i^{(n)}\, h_j^{(n)}\, \Delta^2, \quad h_k^{(n)} = \varphi_{1k}^{(n)}.$$

On the other hand it is well known that the first characteristic value of the boundary problem (7) is equal to the reciprocal of the maximum of the double integral

$$\int_0^1 \int_0^1 K(x, t)\, \varphi(x)\, \varphi(t)\, dx\, dt, \qquad K(x, t) = G(x, t)\, q(x)\, q(t)$$

under the condition

$$\int_0^1 q(x)\, \varphi(x)^2\, dx = 1,$$

and that the corresponding fundamental (normalized) function maximizes this integral, that is, if

$$\frac{d^2 \varphi_1(x)}{dx^2} + \lambda_1\, q(x)\, \varphi(x) = 0, \quad \varphi_1(0) = \varphi_1(1) = 0, \quad \int_0^1 q(x)\, \varphi_1(x)^2\, dx = 1,$$

then

$$\frac{1}{\lambda_1} = \min \int_0^1 \int_0^1 K(x, t)\, \varphi(x)\, \varphi(t)\, dx\, dt = \int_0^1 \int_0^1 K(x, t)\, \varphi_1(x)\, \varphi_1(t)\, dx\, dt.$$

Hence we have

$$\begin{aligned} 0 > \frac{1}{\lambda_{1,n}} - \frac{1}{\lambda_1} &= \sum_{i,j=0}^{n} K_{ij}\, h_i^{(n)}\, h_j^{(n)}\, \Delta^2 - \int_0^1 \int_0^1 K(x, t)\, \varphi_1(x)\, \varphi_1(t)\, dx\, dt \\ &= \left[\sum_{i,j=0}^{n} K_{ij}\, h_i^{(n)}\, h_j^{(n)}\, \Delta^2 - \sum_{i,j=0}^{n} K_{ij}\, h_i\, h_j\, \Delta^2 \right] \\ &\quad + \left[\sum_{i,j=0}^{n} K_{ij}\, h_i\, h_j\, \Delta^2 - \int_0^1 \int_0^1 K(x, t)\, \varphi_1(x)\, \varphi_1(t)\, dx\, dt \right] \end{aligned} \tag{16}$$

whence

$$(17)\quad \left|\frac{1}{\lambda_{1,n}}-\frac{1}{\lambda_1}\right| \leqq \left|\sum_{i,j=0}^{n} K_{ij}\, h_i\, h_j\, \Delta^2 - \int_0^1\int_0^1 K(x,t)\,\varphi_1(x)\,\varphi_1(t)\,dx\,dt\right|,$$

the expression in the first brackets of the right-hand member of (16) being positive. The sequence $h_0, \cdots, h_n$ is arbitrary subject only to (15). It is always possible to find $(n-1)$ points $\xi_1, \xi_2, \cdots, \xi_{n-1}$ such that ξ_i is in (x_{i-1}, x_i) and that

$$\sum_{i=1}^{n-1} q(\xi_i)\,\varphi_1(\xi_i)^2\,\Delta = \int_0^1 q(x)\,\varphi_1(x)^2\,dx.$$

If in (17) we set $h_0 = h = 0$, $h_i = \sqrt{\frac{q(\xi_i)}{q(x_i)}}\,\varphi_1(\xi_i)$, we see at once:

$$(18)\qquad \frac{1}{\lambda_{1,n}}-\frac{1}{\lambda_1} = O\left(\frac{1}{n}\right), \qquad \lambda_{1,n}-\lambda_1 = O\left(\frac{1}{n}\right),$$

because obviously then the right-hand member of (17) is of the order $O\left(\frac{1}{n}\right)$

Now we shall explain a method for estimating the order of the difference $\varphi_1(x_k) - \varphi_{1k}^{(n)}$ which can be applied also to estimating the difference $\varphi_\nu(x_k) - \varphi_{\nu k}^{(n)}$, where $\varphi_\nu(x)$ is the fundamental function of the system (17) corresponding to the characteristic value λ_ν.

We observe first that

$$(19)\qquad \frac{\Delta^2\varphi_1(x_{k-1})}{\Delta x^2} + \lambda_{1,n}\,q(x_k)\,\varphi_1(x_k) = \varepsilon_{k,n}\,q_k \qquad (k = 1, 2, \cdots, n-1)$$

where

$$\varepsilon_{k,n} = \frac{1}{q_k}\left[\frac{\Delta^2\varphi_1(x_{k-1})}{\Delta x^2} - \frac{d^2\varphi_1(x_k)}{dx^2}\right] + (\lambda_{1,n}-\lambda_1)\,\varphi_1(x_k).$$

The function $\frac{d^3\varphi_1(x)}{dx^3}$ being bounded on $(0, 1)$, it follows from (18)

$$(20)\qquad \varepsilon_{k,n} = O\left(\frac{1}{n}\right).$$

Now, using formulas (12), (8_1) we easily obtain an expression for $\varphi_1(x_k)$ considered as a solution of the non-homogeneous equation (19) in finite differences:

$$(21)\quad \varphi_1(x_k) = -\sum_{\nu=2}^{n-1}\frac{\varphi_{\nu,k}^{(n)}\,\varepsilon_\nu}{\lambda_{\nu,n}-\lambda_{1,n}} + C\varphi_{1,k}^{(n)}; \quad \varepsilon_\nu = \sum_{k=0}^{n} q_k\,\varepsilon_{\nu,k}\,\varphi_{\nu,k}^{(n)}\Delta;\ C \text{ const.}$$

We proceed to computing the sum

$$\sum_{k=0}^{n} q_k\,\varphi_1(x_k)^2\,\Delta$$

in two different ways. On the one hand, by virtue of boundary conditions, this sum coincides with a trapezoidal formula for the integral

$$\int_0^1 q(x)\,\varphi_1(x)^2\,dx = 1,$$

so that

$$\sum_{k=0}^{n} q_k\,\varphi_1(x_k)^2\,\Delta = 1+\frac{\eta_n}{n^2};\quad \eta_n \text{ bounded.}$$

On the other hand using (10), (19) and (12) it is readily seen

$$\sum_{k=0}^{n} q_k\,\varphi_1(x_k)^2\,\Delta = 1+\frac{\eta_n}{n^2} = C^2+\sum_{\nu=2}^{n-1}\frac{\varepsilon_\nu^2}{(\lambda_{\nu,n}-\lambda_{1,n})^2} = C^2+O\left(\frac{1}{n^2}\right),$$

whence

$$C = 1+O\left(\frac{1}{n^2}\right).$$

On substituting in (21), we easily find:

$$|\varphi_{1k}^{(n)}-\varphi_1(x_k)| \leqq (C-1)\,|\varphi_{1k}^{(n)}|+\sqrt{\sum_{\nu=2}^{n-1}\frac{(\varphi_{\nu,k}^{(n)})^2}{(\lambda_{\nu,n}-\lambda_{1,n})^2}\sum_{\nu=2}^{n-1}\varepsilon_\nu^2} = O\left(\frac{1}{n}\right).$$

In order to apply the same method to estimating the difference $\varphi_{\nu,k}^{(n)}-\varphi_\nu(x_k)$ we need to know the order of the difference $\lambda_{\nu,n}-\lambda_\nu$. We shall determine this order in the case $\nu = 3$ only; the extension to the general case will not present any essential difficulty.

As before, we have

$$(22)\quad \left|\frac{1}{\lambda_{3,n}}-\frac{1}{\lambda_3}\right| \leqq \left|\sum_{i,j=0}^{n} K_{ij}\,h_i\,h_j\,\Delta^2-\int_0^1\int_0^1 K(x,t)\,\varphi_3(x)\,\varphi_3(t)\,dx\,dt\right|,$$

where $h_0, h_1, \cdots, h_n$ is an arbitrary sequence satisfying the conditions:

$$(23)\quad \sum_{i=0}^{n} q_i\,h_i^2\,\Delta = 1,\quad \sum_{i=0}^{n} q_i\,h_i\,\varphi_{1i}^{(n)}\,\Delta = 0,\quad \sum_{i=0}^{n} q_i\,h_i\,\varphi_{2i}^{(n)}\,\Delta = 0,$$

and $\varphi_{1i}^{(n)}$, $\varphi_{2i}^{(n)}$ are supposed to be known. Let

$$\varphi_{1i}^{(n)}-\varphi_1(x_i) = \zeta_{1i};\quad \varphi_{2i}^{(n)}-\varphi_2(x_i) = \zeta_{2i}.$$

It has been proved already that ζ_{1i}, $\lambda_{1,n}-\lambda_1$ are of the order $O\left(\frac{1}{n}\right)$. We assume that the same is true of ζ_{2i}, $\lambda_{2,n}-\lambda_2$ and we shall prove that a sequence $h_0, h_1, \cdots, h_n$ satisfying all the conditions can be determined so that

$$h_i-\varphi_3(x_i) = O\left(\frac{1}{n}\right).$$

If this is so, then (22) shows immediately that

$$\lambda_{3,n} - \lambda_3 = O\left(\frac{1}{n}\right)$$

and the same reasoning as before will prove that

$$\varphi_{3i}^{(n)} - \varphi_3(x_i) = O\left(\frac{1}{n}\right).$$

We set $h_i - \varphi_3(x_i) = \zeta_i$. If we substitute this in (23) and if we observe that

$$\sum_{i=0}^{n} q_i \varphi_3(x_i)^2 \Delta = 1 + O\left(\frac{1}{n}\right),$$

we have

$$\sum_{i=0}^{n} q_i \{\zeta_i^2 + 2\zeta_i \varphi_3(x_i)\} \Delta = O\left(\frac{1}{n}\right) \tag{24}$$

and in the same way

$$\sum_{i=0}^{n} q_i \zeta_i \varphi_{1i}^{(n)} \Delta = O\left(\frac{1}{n}\right), \qquad \sum_{i=0}^{n} q_i \zeta_i \varphi_{2i}^{(n)} \Delta = O\left(\frac{1}{n}\right), \tag{25}$$

if we take into account the order of ζ_{1i}, ζ_{2i}. Equations (24), (25) for ζ_i are equivalent to (23). We shall try expressions of the form

$$\zeta_i = \frac{\alpha_n}{n}, \text{ if } 0 \leqq i \leqq i_1; \quad \zeta_i = \frac{\beta_n}{n}, \text{ if } i_1 + 1 < i \leqq i_2;$$
$$\zeta_i = \frac{\gamma_n}{n}, \text{ if } i_2 < i \leqq n$$

for the solutions of (24), (25). This gives

$$\begin{cases} \frac{\alpha_n^2}{n^2} Q_1 + \frac{\beta_n^2}{n^2} Q_2 + \frac{\gamma_n^2}{n^2} Q_3 + \frac{\alpha_n}{n} \theta_1 + \frac{\beta_n}{n} \theta_2 + \frac{\gamma_n}{n} \theta_3 = O\left(\frac{1}{n}\right), \\ \frac{\alpha_n}{n} \psi_{11} + \frac{\beta_n}{n} \psi_{12} + \frac{\gamma_n}{n} \psi_{13} = O\left(\frac{1}{n}\right), \\ \frac{\alpha_n}{n} \psi_{21} + \frac{\beta_n}{n} \psi_{22} + \frac{\gamma_n}{n} \psi_{23} = O\left(\frac{1}{n}\right), \end{cases} \tag{26}$$

where

$$Q_1 = \sum_{i=0}^{i_1} q_i \Delta; \qquad Q_2 = \sum_{i=i_1+1}^{i_2} q_i \Delta; \qquad Q_3 = \sum_{i=i_2+1}^{n} q_i \Delta;$$

$$\theta_1 = 2 \sum_{i=0}^{i_1} q_i \varphi_3(x_i) \Delta; \quad \theta_2 = 2 \sum_{i=i_1+1}^{i_2} q_i \varphi_3(x_i) \Delta; \quad \theta_3 = 2 \sum_{i=i_2+1}^{n} q_i \varphi_3(x_i) \Delta;$$

$$\psi_{11} = \sum_{i=0}^{i_1} q_i \varphi_{1i}^{(n)} \Delta; \quad \psi_{12} = \sum_{i=i_1+1}^{i_2} q_i \varphi_{1i}^{(n)} \Delta; \quad \psi_{13} = \sum_{i=i_2+1}^{n} q_i \varphi_{1i}^{(n)} \Delta;$$

$$\psi_{21} = \sum_{i=0}^{i_1} q_i \varphi_{2i}^{(n)} \Delta; \quad \psi_{22} = \sum_{i=i_1+1}^{i_2} q_i \varphi_{2i}^{(n)} \Delta; \quad \psi_{23} = \sum_{i=i_2+1}^{n} q_i \varphi_{2i}^{(n)} \Delta.$$

The system consisting of two last of the equations (26) can be solved with respect to β_n, γ_n and will yield the result of the form

$$\beta_n = A_{1n}\alpha_n + B_{1n}; \qquad \gamma_n = A_{2n}\alpha_n + B_{2n}, \tag{27}$$

where $A_{1n}, A_{2n}, B_{1n}, B_{2n}$ are bounded. This will be proved if we can show that if n is sufficiently large and the indices i_1, i_2 are suitably chosen, the absolute value of the determinant

$$\psi_{12}\psi_{23} - \psi_{13}\psi_{22}$$

has a lower limit >0 as $n \to \infty$. Let be δ_n the upper bound of

$$|\Omega_n(x_{i_1}, x_{i_2})| = \left| \sum_{i_1+1}^{i_2} q_i h_i^{(1)} \Delta \sum_{i_2+1}^{n} q_i h_i^{(2)} \Delta - \sum_{i_2+1}^{n} q_i h_i^{(1)} \Delta \sum_{i_1+1}^{i_2} q_i h_i^{(2)} \Delta \right|$$

in the domain $M\frac{1}{n}$

$$0 \leqq x_{i_1}; \qquad x_{i_1} + \Delta \leqq x_{i_2}; \qquad x_{i_2} + \Delta \leqq 1.$$

It is easy to see that all δ_n are different from zero. Now we will show that the lower bound of these quantities is also different from zero.

Let us suppose the contrary.

Then there exists such a sequence $[V_n]$ that

$$\delta V_n \underset{n\to\infty}{\to} 0.$$

Each point $(\bar{X}_1, \bar{X}_2)$ of the domain M_0

$$0 \leqq X_1; \qquad X \leqq X_2; \qquad X_2 \leqq 1;$$

can be evidently approached by the points $X(i_1)_n$, $X(i_2)_n$ of the domain $M_{\left(\frac{1}{n}\right)}$. Then in virtue of property of $h_i^{(1)}$ and $h_i^{(2)}$ (precedingly demonstrated) we have

$$\Omega(x(i_1)_n, x(i_2)_n) \to \Omega(\bar{X}_1, \bar{X}_2)$$

where

$$\Omega(\bar{X}_1, \bar{X}_2) = \int_{\bar{X}_1}^{\bar{X}_2} q\varphi_1\, dx \cdot \int_{\bar{X}_2}^{1} q\varphi_2\, dx - \int_{\bar{X}_1}^{\bar{X}_2} q\varphi_2\, dx \cdot \int_{\bar{X}_2}^{1} q\varphi_1\, dx$$

and particularly

$$|\Omega(\bar{X}_1, \bar{X}_2)| = \lim_{V_n \to \infty} |\Omega(x(i_1)_{V_n}, x(i_2)_{V_n})| \leqq \lim_{V_n \to \infty} \delta V_n = 0.$$

The point $(\bar{X}_1, \bar{X}_2)$ being arbitrary in M_0, this conclusion is impossible. Therefore the lower limit δ of the sequence δ_n $(n = 1, 2, 3 \cdots)$ is different from zero. Hencefore for each n there exists in $M_{\left(\frac{1}{n}\right)}$ such a point $x(i_1)_n, x(i_2)_n$ that

$$|\Omega(x(i_1)_n, x(i_2)_n)| \geqq \frac{\delta_n}{2} > \frac{\delta}{2} > 0.$$

Now, substitution of (27) in (26) yields an equation of the form

$$\frac{N_n \alpha_n^2}{n} + M_n \alpha_n = C_n,$$

where N_n, M_n, C_n are bounded. For n sufficiently large this equation has a real root which is bounded, which proves that $\alpha_n, \beta_n, \gamma_n$ are bounded. Thus the relations

$$\lambda_{3,n} - \lambda_3 = O\left(\frac{1}{n}\right), \qquad \varphi_{3,k}^{(n)} - \varphi_3(x_k) = O\left(\frac{1}{n}\right)$$

are established.

4. **Rayleigh's Principle from the point of view of a minimum problem.** The system in finite differences:

$$\begin{aligned} \frac{\Delta^2 y_{k-1}^{(n)}}{\Delta x^2} &= q_k y_k^{(n)} + f_k \qquad (k = 1, 2, \cdots, n-1) \\ y_0^{(n)} &= y_n^{(n)} = 0 \end{aligned} \tag{2}$$

can be considered as Euler's equations corresponding to the minimum of the sum

$$\begin{aligned} S_n(y) &= \sum_{i=0}^{n-1} \left[\left(\frac{\Delta y_i}{\Delta x}\right)^2 + q_i y_i^2 + 2 f_i y_i\right] \Delta \\ &= \sum_{i=0}^{n-1} \left[\left(\frac{\Delta y_i}{\Delta x}\right)^2 + q_i \left(y_i + \frac{f_i}{q_i}\right)^2\right] \Delta + \text{Const.} \end{aligned}$$

which shows at once that the solution of (2) is unique.

Let $y = y^{(n)}(x)$ be the equation of the polygon whose vertices are at the points $(x_k, y_k^{(n)})$. We set

$$I(y) = \int_0^1 \left[\left(\frac{dy}{dx}\right)^2 + q(x)\, y(x)^2 + 2 f(x)\, y(x)\right] dx,$$

$$\bar{S}_n(y) = \sum_{i=0}^{n-1}\left[\left(\frac{\Delta y_i}{\Delta x}\right)^2 + q_i\left(y_i + \frac{f_i}{q_i}\right)^2\right]\Delta$$

and we proceed to prove that

$$I(y^{(n)}) - S_n(y^{(n)}) = \sum_{i=0}^{n-1}\left\{\int_{x_i}^{x_i+\Delta}[q(x)\,y^{(n)}(x)^2 + 2f(x)\,y^{(n)}(x)]\,dx - [q_i\,y_i^{(n)2} + 2f_i\,y_i^{(n)}]\,\Delta\right\} = O\left(\frac{1}{n}\right).$$

This will be proved immediately, if we show that $\Delta y_i^{(n)}/\Delta x$ is bounded. This is certainly true of the sum $S_n(y^{(n)})$, whence the sum

$$\sum_{i=0}^{n-1}\left[\frac{\Delta y_i^{(n)}}{\Delta x}\right]^2\Delta$$

and therefore $y_i^{(n)}$ is bounded. Equation (2) shows then that $\Delta^2 y_i^{(n)}/\Delta x^2$ is also bounded, which implies the boundedness of $\Delta y_i^{(n)}/\Delta x$.

From the Calculus of Variations it is well known that the problem of minimizing the integral $I(y)$ is regular and admits of a unique solution, because of the condition $q(x) > 0$. If this solution is $y(x)$, then $y(x)$ is also the solution of the differential system (1). Now, observing the extremal properties of the functions $y(x)$, $y^{(n)}(x)$, we have

$$0 < I(y^{(n)}) - I(y) = [I(y^{(n)}) - S_n(y^{(n)})] + [S_n^{(n)}(y^{(n)}) - S_n(y)] + [S_n(y) - I(y)].$$

The middle term of the right-hand member being negative and two other terms being of the order $O\left(\frac{1}{n}\right)$, this shows

$$I(y^{(n)}) - I(y) = O\left(\frac{1}{n}\right).$$

Using this result, a simple application of Taylor's formula will enable us to estimate the order of the difference $y(x) - y^{(n)}(x)$, after which equations (1) and (2) will show

$$\lim \frac{\Delta^2 y_{i-1}^{(n)}}{\Delta x^2} = \frac{d^2 y(x)}{dx^2}, \quad \text{as} \quad x_i = \frac{i}{n} \to x$$

and will give an estimate of the order of the difference $\dfrac{\Delta^2 y_{i-1}^{(n)}}{\Delta x^2} - \dfrac{d^2 y}{dx^2}$. In analogous way we may justify Rayleigh's Principle in the case $q(x) < 0$, where it can be applied to the computation of characteristic values. Consider for instance the problem of minimizing the sum

$$\Sigma_n(y) = \sum_{i=0}^{n-1}\left[\left(\frac{\Delta y_i}{\Delta x}\right)^2 - q_i\, y_i^2\right]\Delta^*$$

under the conditions:

$$\sum_{i=0}^{n} q_i\, y_i^2\, \Delta = 1, \qquad y_0 = y_n = 0. \tag{28}$$

It is readily seen that the minimum in question exists and that the minimizing sequence $y_0^{(n)}, \cdots, y_n^{(n)}$ is uniquely determined and represents the solution of the system in finite differences:

$$\frac{\Delta^2 y_{k-1}^{(n)}}{\Delta x^2} + \lambda_{1,n}\, q_k\, y_k^{(n)} = 0; \qquad \sum_{k=0}^{n} q_k\, y_k^{(n)} = 1, \qquad y_0^{(n)} = y_k^{(n)} = 0 \tag{29}$$

the minimum value of the sum $S_n(y)$ being equal to $S_n(y^{(n)}) = \lambda_{1,n}$. On the other hand the function $\varphi_1(x)$ minimizes the integral

$$I(y) = \int_0^1 \left[\left(\frac{dy}{dx}\right)^2 - q(x)\, y(x)^2\right] dx$$

under the condition

$$\int_0^1 q(x)\, y(x)^2\, dx = 1,$$

and the minimum value of the integral $I(y)$ is $I(\varphi_1) = \lambda_1$. Now we see

$$|\lambda_{1,n} - \lambda_1| \leqq |S_n(h) - I(\varphi_1)|$$

where $h_0, h_1, \cdots, h_n$ is an arbitrary sequence subject to (28). By choosing

$$h_0 = h_n = 0; \qquad h_i = \sqrt{\frac{q(\xi_i)}{q(x_i)}}\,\varphi_1(\xi_i) \quad (i = 1, 2, \cdots, n-1);$$

$$\sum_{i=1}^{n-1} q(\xi_i)\,\varphi_1(\xi_i)^2\,\Delta = 1$$

we can fix the order of the difference $\lambda_{1,n} - \lambda_1$ and then we can proceed as in § 3.

5. **The simplest problem of the Calculus of Variations from the point of view of Rayleigh's Principle.** In one of his earliest papers on the Calculus of Variations Euler considered a finite sum instead of an integral, thus obtaining a heuristic method for deducing the fundamental equations of the Calculus of Variations which now are known as

* $q(x)$ has the same meaning as in § 1.

Euler's equations. In the development of the Calculus of Variations, as we are aware, this point of view was abandoned until a recent paper by R. B. Robbins,* who has treated of the same problem under rather general conditions as to the function under the sign of integral. Robbins' method, however, does not seem to enable one to estimate the order of the error at the nth step of approximation. This can be done by using the following method which requires, it is true, more restrictions concerning the function $f\left(x, y, \frac{dy}{dx}\right)$ under the sign of integral:

$$I(y) = \int_0^1 f\left(x, y, \frac{dy}{dx}\right) dx.$$

We suppose namely that

$$(30) \quad \frac{\partial^2 f}{\partial x^2} \geqq \alpha > 0; \quad \left[\frac{\partial^2 f}{\partial y \partial p}\right]^2 - \frac{\partial^2 f}{\partial y^2}\frac{\partial^2 f}{\partial^2 p} \leqq -\beta < 0; \quad y' = p.$$

Then we can easily prove the existence of the absolute minimum of the sum

$$S_n(y) = \sum_{i=0}^{n-1} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right)\Delta.$$

In order to do this we observe that

$$(31) \quad \begin{aligned} f(x, y, p) = {} & f(x, 0, 0) + y f'_y(x, 0, 0) + p f'_p(x, 0, 0) \\ & + \frac{1}{2}\left[\left(\frac{\partial^2 f}{\partial^2 y}\right) y^2 + 2\left(\frac{\partial^2 f}{\partial y \partial p}\right) p y + \left(\frac{\partial^2 f}{\partial p^2}\right) p^2\right] \end{aligned}$$

where

$$\left(\frac{\partial^2 f}{\partial y^2}\right) = f''_{y^2}(x, \theta y, \theta p); \quad \left(\frac{\partial^2 f}{\partial y \partial p}\right) = f''_{yp}(x, \theta y, \theta p);$$

$$\left(\frac{\partial^2 f}{\partial p^2}\right) = f''_{p^2}(x, \theta y, \theta p), \qquad 0 < \theta < 1.$$

In virtue of the assumptions made the expression in brackets is a definite positive form in y and p, such that:

$$(32) \quad \left(\frac{\partial^2 f}{\partial y^2}\right) y^2 + 2\left(\frac{\partial^2 f}{\partial y \partial p}\right) yp + \left(\frac{\partial^2 f}{\partial p^2}\right) p^2 \geqq \alpha(y^2 + p^2).$$

On setting

$$k_1 = \frac{1}{2\alpha} f'_y(x, 0, 0); \qquad k_2 = \frac{1}{2\alpha} f'_p(x, 0, 0),$$

* Loc. cit.

we obviously have

$$f(x, y, p) - C(x) \geqq \alpha[(y+k_1)^2+(p+k_2)^2], \tag{33}$$

where $C(x)$ is a function in x alone. From (33) it is readily seen

$$S_n(y-k_1) - \sum_{i=0}^{n-1} C(x_i)\Delta \geqq \alpha \sum_{i=0}^{n-1} y_i^2 \Delta. \tag{34}$$

Now, if we consider the sequence $y_0^{(m)}, y_1^{(m)}, \cdots, y_n^{(m)}$ as representing a point $Y^{(m)}$ in $(n+1)$-space, and if $Y^{(1)}, \cdots, Y^{(m)}, \cdots$ is a sequence of points for which $S_n(y)$ decreases and approaches its lower bound, it is obvious that the points $Y^{(m)}$ are within a hypersphere of the radius $\sqrt{\frac{K+C}{\alpha\Delta}}$ where K and C are some upper bounds of $S_n(y^{(m)}-k_1)$ and $|C(x)|$ respectively. The sequence $Y^{(1)}, Y^{(2)}, \cdots, Y^{(m)}, \cdots$ then has at least one limiting point within the same hypersphere, which, because of the continuity of $S_n(y)$ with respect to $y_0, y_1, \cdots, y_n$, furnishes the absolute minimum of the sum $S_n(y)$. The minimizing sequence $y_0^{(n)}, \cdots, y_n^{(n)}$ thus obtained, is unique, because, if $0 = y_0, y_1, \cdots, y_n = 0$ is any other sequence, we have; setting $\eta_i = y_i = y_i^{(n)}$

$$\begin{aligned} S_n(y) - S_n(y^{(n)}) = {} & \sum_{i=0}^{n-1} \frac{\partial S_n(y)}{\partial y}\bigg|_{y=y^{(n)}} \eta_i^{(n)}\Delta + \frac{1}{2}\sum_{i=0}^{n-1}\Delta\left\{\left(\frac{\partial^2 f}{\partial y^2}\right)\eta_i^{(n)2}\right. \\ & \left. + 2\left(\frac{\partial^2 f}{\partial y\,\partial p}\right)\eta_i^{(n)}\frac{\Delta\eta_i^{(n)}}{\Delta x} + \left(\frac{\partial^2 f}{\partial p^2}\right)\left(\frac{\Delta\eta_i^{(n)}}{\Delta y}\right)^2\right\}. \end{aligned} \tag{35}$$

The first term of the right-hand member is zero, as this follows from the extremal property of the sequence $y_0^{(n)}, y_1^{(n)}, \cdots, y_n^{(n)}$. The second term is positive, unless

$$\eta_i = 0, \text{ i. e., } y_i = y_i^{(n)} \qquad (i = 0, 1, \cdots, n),$$

in virtue of (30). Thus

$$S_n(y) > S_n(y^{(n)}),$$

unless the sequences y_k and $y_k^{(n)}$ are identical.

The minimizing sequence $y_0^{(n)}, \cdots, y_n^{(n)}$ satisfies the system in finite differences:

$$\begin{aligned} & f_p'(x_k, y_k^{(n)}, p_k^{(n)}) - f_p'(x_{k-1}, y_{k-1}^{(n)}, p_{k-1}^{(n)}) - \Delta f_y'(x_k, y_k^{(n)}, p_k^{(n)}) = 0, \\ & y_0^{(n)} = y_n^{(n)} = 0, \qquad p_k^{(n)} = \frac{\Delta y_k^{(n)}}{\Delta x} = \frac{y_{k+1}^{(n)} - y_k^{(n)}}{\Delta x}, \\ & \qquad\qquad (k = 1, 2, \cdots, n-1) \end{aligned} \tag{36}$$

which consequently has a unique solution.

Let $y(x)$ be the function which minimizes the integral

$$I(y) = \int_0^1 f\left(x, y, \frac{dy}{dx}\right) dx, \qquad y(0) = y(1) = 0.$$

The corresponding problem of the Calculus of Variations being regular, $y(x)$ is a (unique) solution of Euler's differential system

$$\frac{d}{dx} f'_p(x, y, y') - f'_y(x, y, p) = 0; \qquad y(0) = y(1) = 0. \tag{37}$$

Because of the extremal properties of the sequence $y_k^{(n)}$ and the function $y(x)$, we have

$$\begin{aligned} 0 < I(y^{(n)}) - I(y) &= [I(y^{(n)}) - S_n(y^{(n)})] + [S_n(y^{(n)}) - S_n(y)] \\ &+ [S_n(y) - I(y)] < [I(y^{(n)} - S_n(y^{(n)})] + [S_n(y) - I(y)], \end{aligned} \tag{38}$$

where $y = y^{(n)}(x)$ is the equation of the polygon with the vertices at the points $(x_k, y_k^{(n)})$. But

$$I(y) - S_n(y) = O\left(\frac{1}{n}\right). \tag{39}$$

On the other hand we have

$$\begin{aligned} |I(y^{(n)}) - S_n(y^{(n)})| &= \left| \sum_{i=0}^{n-1} \left\{ f\left(x_i, y_i^{(n)}, \frac{\Delta y_i^{(n)}}{\Delta x}\right) \Delta - \int_{x_i}^{x_i + \Delta} f\left(x, y, \frac{\Delta y_i^{(n)}}{\Delta x}\right) dx \right\} \right| \\ &\leqq \Delta \left\{ \max \left| \frac{\partial f}{\partial x} \right| + \max \left| \frac{\partial f}{\partial y} \right| \cdot \left| \frac{\Delta y_i^{(n)}}{\Delta x} \right| \right\}. \end{aligned} \tag{40}$$

The sums $S_n(y^{(n)})$ being bounded (with respect to n), from (33) it is readily seen that the sum

$$\sum_{i=0}^{n-1} \left[\frac{\Delta y_i^{(n)}}{\Delta x} \right]^2 \Delta$$

is also bounded, whence $y_k^{(n)}$ is bounded. Now, system (36) may be rewritten as follows:

$$(f''_{p^2}) \frac{\Delta^2 y_{k-1}^{(n)}}{\Delta x^2} = (f'_y) - (f''_{xp}) - (f''_{yp}) \frac{\Delta y_{k-1}^{(n)}}{\Delta x}.$$

If we assume that

$$|f''_{yp}| < M, \quad |f'_y| \text{ and } |f''_{xp}| < N|y| + P|p|, \tag{41}$$

where M, N, P are positive constants, we see at once that

$$\sum_{i=1}^{n-1}\left[\frac{\Delta^2 y_{i-1}^{(n)}}{\Delta x^2}\right]^2 \Delta$$

is bounded. It is easy to prove then that $\Delta y_{k,n}^{(n)}/\Delta x$ is bounded. Using (38), (39) and (40) we see that

$$I(y^{(n)}) - I(y) = O\left(\frac{1}{n}\right).$$

Now, an argument analogous to that which has been used in proving the uniqueness of the minimizing sequence $y_0^{(n)}, \cdots, y_n^{(n)}$ will show us that

$$\int_0^1 \left[\frac{d(y-y^{(n)})}{dx}\right]^2 dx = O\left(\frac{1}{n}\right),$$

whence, finally: $y(x) - y^{(n)}(x) = O\left(\frac{1}{n}\right)$. Thus the following theorem is established:

THEOREM. *If the function $f(x, y, p)$ of the simplest problem of the Calculus of Variations satisfies conditions* (30) *and* (41) *the application of Rayleigh's Principle to this problem is justified and the order of the error at the n-th approximation can be estimated.*

6. **Application of the method of least squares to the approximate solution of the equations in finite differences.** Using the basic idea of the method of least squares* we shall try an approximate solution of the system (2) in form of a linear aggregate

$$y_{i,n}^{(k)} = \sum_{j=1}^{k} a_j \psi_i^{(j)} \qquad (k \leqq n-1)$$

which minimizes the sum

$$\sigma_n(y_i) = \sum_{i=1}^{n-1}\left[\frac{\Delta^2 y_{k-1}}{\Delta x^2} - q_i y_i - f_i\right]^2 \Delta$$

and where $\psi_i^{(j)}$ are solutions of systems in finite differences:

$$\frac{\Delta^2 u_{i-1}}{\Delta x^2} + \lambda_j u_i = 0, \quad u_0 = u_n = 0 \qquad (i = 1, 2, \cdots, n-1), \tag{42}$$

i. e.

$$\psi_i^{(j)} = \sqrt{2} \sin \frac{ij\pi}{n} \quad (j = 1, 2, \cdots, k;\ i = 0, 1, \cdots, n).$$

* Cf. N. Kryloff, Sur une méthode basée sur le principe de minimum pour la solution approchée des équations differentielles, Comptes Rendus Paris, vol. 181, p. 86, where the same method has been discussed in application to the differential boundary problems.

The conditions of minimum yield a system of k linear equations in k unknowns $a_1, a_2, \cdots, a_k$, which certainly admits of a unique solution if $q_k > 0$. Let

$$Y_{i,n}^{(k)} = \sum_{j=1}^{k} A_j \psi_i^{(j)}$$

denote the trigonometric polynomial which interpolates the solution $y_i^{(n)}$ of (4). Using the extremal property of the sequence $y_0^{(n)}, \cdots, y_n^{(n)}$ it is seen that

$$|\sigma_n(y_{i,n}^{(k)}) - \sigma_n(y_i^{(n)})| \leqq |\sigma_n(Y_{i,n}^{(k)}) - \sigma_n(y_i^{(n)})| = \varepsilon_{k,n}, \tag{43}$$

where, as it may be easily proved,

$$\varepsilon_{k,n} = \sum_{i=1}^{n-1} \left[\frac{\Delta^2 (Y_{i-1,n}^{(k)} - y_{i-1}^{(n)})}{\Delta x^2} - q_i (Y_{i,n}^{(k)} - y_i^{(n)}) \right]^2 \Delta.$$

Hence

$$\begin{aligned} \varepsilon_{k,n} &\leqq 2 \sum_{i=1}^{n-1} \left[\frac{\Delta^2 (Y_{i-1,n}^{(k)} - y_{i-1}^{(n)})}{\Delta x^2} \right]^2 \Delta + 2 \max q_i^2 \sum_{i=1}^{n-1} [Y_{i,n}^{(k)} - y_i^{(n)}]^2 \Delta \\ &\leqq M \sum_{i=1}^{n-1} \left[\frac{\Delta^2 (Y_{i-1,n}^{(k)} - y_{i-1}^{(n)})}{\Delta x^2} \right] \Delta; \end{aligned} \tag{44}$$

$$M = 2 + \frac{1}{4} \max q_i^2,$$

which can be easily proved by using well known properties of the Green's function $G_{i,k}$. On setting

$$[\omega_i]_k = \sum_{j=1}^{k} \psi_i^{(j)} \sum_{l=0}^{n} \omega_l \psi_l^{(j)} \Delta,$$

we have

$$\frac{\Delta^2 (Y_{i-1,n}^{(k)} - y_{i-1}^{(n)})}{\Delta x^2} = \left[\frac{\Delta^2 y_{i-1}^{(n)}}{\Delta x^2} \right]_k - \frac{\Delta^2 y_{i-1}^{(n)}}{\Delta x^2} = \varepsilon_{k,n}$$

whence

$$\varepsilon_{k,n} \leqq M \sum_{j=k+1}^{n-1} \left\{ \sum_{l=0}^{n} \frac{\Delta^2 y_l^{(n)}}{\Delta x^2} \psi_l^{(j)} \Delta \right\} \leqq M \frac{n-k}{n} \frac{1}{k} \sum_{l=0}^{n} \left[\frac{\Delta^3 y_l^{(n)}}{\Delta x^3} \right]^2 \Delta, \tag{45}$$

because

$$\left| \sum_{l=0}^{n} \omega_l \psi_l^{(j)} \Delta \right| \leqq \frac{1}{j} \sqrt{\sum_{l=0}^{n} \left[\frac{\Delta \omega_l}{\Delta x} \right]^2 \Delta}; \quad \sum_{j=k+1}^{n-1} \frac{1}{j^2} < \frac{n-k}{n} \frac{1}{k}.$$

If we assume that $q(x)$ and $f(x)$ satisfy a Lipschitz's condition, we obtain from (45)

$$\varepsilon_{k,n} < M\frac{n-k}{nk}L,$$

where L is a positive constant. Now (43) shows

$$\sum_{i=0}^{n}\left[\frac{\Delta^2(y_{i-1,n}^{(k)} - y_{i-1}^{(n)})}{\Delta x^2} - q_i(y_{i,n}^{(k)} - y_i^{(n)})\right]^2\Delta$$

and then a reasoning analogous to that which has been used by the Author in the case of the differential boundary problem, will yield an estimate of the form

$$|y_{i,n}^{(k)} - y_i^{(n)}| < A\sqrt{\varepsilon_{k,n}} < A\sqrt{ML}\sqrt{\frac{n-k}{nk}},$$

where A is a positive constant.

It has been proved in § 1 that

$$y_i^{(n)} \to y(x), \text{ as } \frac{i}{n} \to x; \quad \frac{d^2y(x)}{dx^2} - q(x)\,y(x) = f(x); \quad y(0) = y(1) = 0.$$

Hence the method above may be considered as a method for an approximate solution of the differential boundary problem which corresponds to system (2) in finite differences. This method may present some advantages in the case when n is so large that an immediate algebraic solution of the system (2) becomes tiresome. The order of the error can be easily evaluated.

7. **Boundary value problems in two dimensions.** For the sake of brevity we shall treat only of the simplest elliptic case, namely

$$\frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2} - q(x,y)\,u(x,y) = f(x,y), \tag{46}$$
$$u(x,y) = 0 \text{ on the contour } (C).$$

We assume that the contour (C) is a square with vertices at the points $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$ and that $q(x,y) > 0$ in the interior of (C). It seems probable, however, that the method below can be extended to more general cases, for instance, to those which have been treated in a recent paper by M. Plancherel* on Ritz's method.

The corresponding system in finite differences is now

$$\frac{\Delta^2 u_{i-1\,j-1}^{(n)}}{\Delta x^2} + \frac{\Delta^2 u_{i-1\,j-1}^{(n)}}{\Delta x^2} - q_{ij}\,u_{ij}^{(n)} = f_{ij}; \quad \Delta x = \Delta y = \frac{1}{n}, \tag{47}$$
$$u_{ij}^{(n)} = O \text{ on } (C).$$

* Sur la méthode d'intégration de Ritz, Bull. des Sciences math., (2), vols. 47, 48 (1923–1924).

System (47) admits of a unique solution, because of the condition $q_{ij} \geqq 0$. On setting

$$\delta_{ij} = u_{ij}^{(n)} - u_{ij}$$

we have

$$(48) \qquad \frac{\Delta^2 \delta_{i-1\,j-1}}{\Delta x^2} + \frac{\Delta^2 \delta_{i-1\,j-1}}{\Delta y^2} - q_{ij}\,\delta_{ij} = \varepsilon_{ij}^{(n)}, \quad \delta_{ij} = 0 \text{ on } (C),$$

where

$$\varepsilon_{ij}^{(n)} = \frac{\partial^2 u(x_i, y_j)}{\partial x^2} + \frac{\partial^2 u(x_i, y_j)}{\partial y^2} - \frac{\Delta^2 u_{i-1\,j-1}}{\Delta x^2} - \frac{\Delta^2 u_{i-1\,j-1}}{\Delta y^2} = O\left(\frac{1}{n}\right)$$

if we assume that $q(x, y)$ and $f(x, y)$ possess derivatives which are bounded inside the contour (C).

Multiplying (48) by $\delta_{ij}\,\Delta^2$ and summing by parts, we get:

$$\sum_{i,j} \left\{ \left(\frac{\Delta \delta_{ij}}{\Delta x}\right)^2 + \left(\frac{\Delta \delta_{ij}}{\Delta y}\right)^2 + q_{ij}\,\delta_{ij}^2 + \varepsilon_{ij}\,\delta_{ij} \right\} \Delta^2 = 0.$$

Hence, by Cauchy's inequality,

$$(49) \qquad \sum_{i,j} \delta_{ij}^2\,\Delta^2 = O\left(\frac{1}{n}\right), \quad \sum_{i,j} \left(\frac{\Delta \delta_{ij}}{\Delta x}\right)^2 \Delta^2 = O\left(\frac{1}{n}\right), \quad \sum_{i,j} \left(\frac{\Delta \delta_{ij}}{\Delta x}\right)^2 \Delta^2 = O\left(\frac{1}{n}\right).$$

Now we can apply, mutatis mutandis, the method used by Prof. Richardson* in establishing the existence of fundamental solutions of the two-dimensional problem (without estimating the order of the error, however). For this purpose we set

$$v_{ij} = \frac{\Delta \delta_{ij}}{\Delta x}$$

and we observe that

$$(50) \qquad \frac{\Delta^2 v_{i-1\,j-1}}{\Delta x^2} + \frac{\Delta^2 v_{i-1\,j-1}}{\Delta y^2} = q_{ij}\,v_{ij} + \frac{\Delta q_{ij}}{\Delta x}\,\delta_{ij} + \frac{\Delta \varepsilon_{ij}^{(n)}}{\Delta x},$$

$$(51) \qquad v_{i0} = v_{in} = 0, \quad \frac{\Delta v_{i0}}{\Delta x} = \frac{\Delta v_{in}}{\Delta x} = 0, \quad \frac{\Delta v_{0j}}{\Delta x} = \varepsilon_{0j}, \quad \frac{\Delta v_{nj}}{\Delta x} = \varepsilon_{nj}.$$

Multiplying (50) by $v_{ij}\,\Delta^2$ and summing by parts, we obtain

$$(52) \qquad \sum_{i,j} \left\{ \left[\frac{\Delta v_{ij}}{\Delta x}\right]^2 + \left[\frac{\Delta v_{ij}}{\Delta y}\right]^2 \right\} \Delta^2 = -\sum_{i,j} \left[q_{ij}\,v_{ij}^2 + \frac{\Delta q_{ij}}{\Delta x}\,v_{ij}\,\delta_{ij} + \frac{\Delta \varepsilon_{ij}^{(n)}}{\Delta x}\,v_{ij} \right] \Delta^2 + \sum_{(C)} \left[\frac{\Delta v_{ij}}{\Delta x}\right] v_{ij}\,\Delta - \sum_{(C)} \left[\frac{\Delta v_{ij}}{\Delta y}\right] v_{ij}\,\Delta.$$

* Transactions of the Amer. Math. Soc., vol. XVIII, no. 4.

If we observe that, in virtue of (51),

$$\sum_{(C)}\left[\frac{\Delta v_{ij}}{\Delta x}\right]v_{ij}\Delta = 0, \quad \sum_{(C)}\left[\frac{\Delta v_{ij}}{\Delta y}\right]v_{ij}\Delta = O\left(\frac{1}{n}\right)$$

and if we take into account (52), an easy application of Cauchy's inequality will show:

$$\sum_{i,j}\left\{\left[\frac{\Delta v_{ij}}{\Delta x}\right]^2+\left[\frac{\Delta v_{ij}}{\Delta y}\right]^2\right\}\Delta^2 = O\left(\frac{1}{n}\right),$$

under the assumption *that* $q(x, y)$ *and* $f(x, y)$ *possess bounded derivatives of the third order.*

Then a second application of Cauchy's inequality and boundary conditions will give us:

$$|\delta_{ij}| \leqq \sqrt{\sum_{i,j}\left[\frac{\Delta^2\delta_{ij}}{\Delta x\,\Delta y}\right]^2\Delta^2} \leqq \sqrt{\sum_{i,j}\left\{\left[\frac{\Delta v_{ij}}{\Delta x}\right]^2+\left[\frac{\Delta v_{ij}}{\Delta y}\right]^2\right\}\Delta^2} = O\left(\frac{1}{n}\right),$$

whence

$$u_{ij}^{(n)} - u_{ij} = O\left(\frac{1}{n}\right).$$

This estimate certainly can be improved under further restrictions upon $q(x, y)$ and $f(x, y)$. It may be expected also that the results above can be extended to the case $q<0$ and to the computation of fundamental functions, as well as to the case of a contour (C) different from a square, by using a suitable application of Mr. Richardson's inequalities.

ON VARIOUS CONCEPTIONS OF CORRELATION.*

By Frank M. Weida.

Introduction. The theory of correlation was first developed on definite assumptions as to the form of the destribution of frequency, the so-called normal distribution being assumed. It seems that Sir Francis Galton was the first to deal with correlation among direct observations. He† developed the practical method, determining his coefficient—Galton's function—graphically. Bravais‡ introduced a product-sum, but not a single symbol for a coefficient of correlation. Edgeworth§ developed the theoretical side further on the assumption of a normal distribution. Pearson‖ introduced the product-sum formula on the assumption of a normal distribution. Yule¶ developed the regression method of approaching correlation.

In the problem of correlation, we wish to find some measure of agreement or disagreement between two sets of paired characters. Under the term *correlation* two very different meanings have frequently been considered, *connection* and *concordance.* Gini** established a distinction between the different meanings of connection and concordance. His researches deal with

* Received April 14, in revised form November 28, 1927. Presented by request of the Program Committee at the meeting of the Mathematical Association of America held in Philadelphia, December 30, 1926.

† Galton, Francis, "Regression towards mediocrity in hereditary stature", Jour. Anthrop. Inst., vol. 15 (1886), p. 246; "Family likeness in stature", Proc. Roy. Soc., vol. 40 (1886), p. 42; "Correlations and their measurement", Proc. Roy. Soc., vol. 45 (1888), p. 135.

‡ Bravais, A., "Analyse mathématique sur les probabilités des erreurs de situation d'un point", Acad. des Sciences: Mémoires présentés par divers savants, IIe serie, t. 9 (1846), p. 225.

§ Edgeworth, F. Y., "On correlated averages", Phil. Mag., (5), vol. 34, (1892), p. 190.

‖ Pearson, Karl, "Regression, heredity, and panmixia", Phil. Trans. Roy. Soc., ser. A, vol. 187 (1896), p. 253.

¶ Yule, G. U., "On the significance of Bravais' formulae for regression, etc., in the case of skew correlation", Proc. Roy. Soc., vol. 60 (1897), p. 477; "On the theory of correlation", Jour. Roy. Stat. Soc., vol. 60 (1897), p. 812.

** Gini, C.: "Variabilità e Mutabilità"—Bologna Cuppini 1912; "Sulla misura della concentrazione e della variabilità dei caratteri"—Atti del R. Instituto Veneto di S. L. A., 1913–14, T. 73, P. II; "Di una misura della dissomiglianza tra due gruppi di quantità e della sua applicazione allo studio delle relazioni statistiche", id. 1914–15, T. 74, P. II; "Indici di omofilia e di rassomiglianza e loro relazioni col coefficiente di correlazione e con gli indici di attrazione", id. id. 1914–1915, T. 74, P. II; "Nuovi contributi alla teoria delle relazioni statistiche", id. id. 1914–1915, T. 74, P. II; "Sul criterio di concordanca tra due caratteri", id. id. 1915–16, T. 75, P. II; "Indici di concordanza", id. id. 1915–16, T. 75, P. II; "Delle relazioni tra le intensita congraduate di due caratteri", id. id. 1916–17, T. 76, P. II; "Di una estensione del concetto di scostamento medio etc."; id. id. 1917–18, T. 77, P. II.

the quantitative or qualitative modalities of two characters which are distributed in conformity with rectilinear or disconnected series. He introduced, first, a new concept of dissimilarity between two statistical distributions and put it as the basis of the theory of statistical relations; and secondly, he established a suitable measure for each of the different aspects according to which the quantitative or qualitative modalities of two characters may be presented. As a result of Gini's research, Pietra* established a theory of dissimilarity, connection, and concordance between cyclical series analogous to that of Gini's for rectilinear series.

The theory of correlation that has been much applied in recent years to statistical data has been developed largely as an extension of error theory.† The theory of correlation may also be developed as an integral part of the theory of a-priori probability. Rietz‡ used certain urn schemata as a basis for the development of the theory of correlation. Tschuprow§ has made correlation theory an integral part of the theory of probability and has approached the whole subject from an a-priori point of view. He, first, distinguishes with extreme precision between the case of a functional relation between two variables and of a stochastic (or probability) relation; secondly, he describes the more important measures of correlation; and thirdly, he discusses the problem of estimating from the sample the character of the universe.

The material of this paper is divided into three parts:

Part I deals with various concepts of correlation developed largely as an extension of error theory.|| These concepts have been developed by the English school of statisticians, with Karl Pearson as the most prominent contributor.

Part II is devoted to the theory of connection and concordance. These concepts have been developed by Italians with Gini as the most prominent contributor.

Part III deals with the theory of correlation developed as an integral part of the theory of a-priori probability. These concepts have been developed by Russians with Tschuprow as the most outstanding contributor.

---

* Pietra, G., "The theory of statistical relations with special reference to cyclical series", Metron, vol. 4 (1924–25), pp. 383–557.

† Bravais, "Analyse Mathématique sur les Probabilités des Erreurs de Situation d'un Point", Memoirs par divers Savants, 1846.

‡ Rietz, H. L., "Urn schemata as a basis for the development of correlation theory", Ann. of Math., vol. 21 (1920), pp. 307–322.

§ Tschuprow, A. A., "Grundbegriffe und Grundprobleme der Korrelationstheorie", Leipzig-Berlin, (1925).

|| Rietz, H. L., and Others: "Handbook of Mathematical Statistics" (1924), pp. 120–150; "On certain topics in the mathematical theory of statistics", by Rietz, Bull. Amer. Math. Soc., vol. 30, pp. 429–439.

PART I.

Correlation.*

1. Let us assume data consisting of pairs corresponding values (x_1, y_1), $(x_2, y_2), \cdots, (x_n, y_n)$ where we are interested in a quantitative description of the association of the x's and the corresponding y's. If such a set of pairs of numbers is represented by a system of dots, marking the rectangular coördinates of points, we obtain a so-called "scatter-diagram". From the graph of such data we see that, with an assigned value of x, the corresponding value of y may have many values and thus can not be accurately predicted by the use of a single-valued function of x. The x's and y's are not independent in the probability sense of independence. There is a tendency for the dots of the scatter-diagram to fall into a sort of band which can be fairly well described.

The theory of correlation is devoted to the description and characterization of the field of association between perfect dependence given by a single-valued mathematical function at the one extreme and perfect independence in the probability sense at the other extreme.

2. **Definitions of correlation.** (a) According to Sir Francis Galton,† "Two variable organs are said to be co-related, when the variation of the one is accompanied on the average by more or less variation of the other, and in the same direction." (b) According to Karl Person,‡ "Two organs in the same individual or in a connected pair of individuals are said to be correlated, when a series of the first organ of a definite size being selected, the mean of the sizes of the corresponding second organ is found to be a function of the size of the selected first organ. If the mean is independent of this size, the organs are said to be non-correlated". (c) According to Karl Pearson,§ "Two variables or characters A and B are said to be correlated when, with different values x of A, we do not find the same value y of B equally likely to be associated. In other words, certain values of B are relatively more likely to occur with a value x of A than others. The distribution of B's associated with a given value x of A is termed an x-array of B's. If N pairs of A and B are taken, and n_x of these have the charater $A = x$, these n_x form the x-array of B's. This array, like any other frequency distribution, will have its mean, which we will denote by $\overline{y}_x$, and its standard deviation, which we will denote by σ_{n_x}. The mean of all the B characters shall be $\overline{y}$ and their variability given

* Rietz, H. L., and Others: "Handbook of Mathematical Statistics", (1924), pp. 120–150.

† Proc. Roy. Soc., vol. 45 (1889), p. 135.

‡ Phil. Trans. Roy. Soc., A, vol. 187 (1895), p. 233.

§ Pearson, Karl, "On the general theory of skew correlation and non-linear regression", Drapers' Company Research Memoirs, Biometric Series II (1904), pp. 9–10.

by the standard deviation σ_y. Similarly $\bar{x}$, σ_x, will denote the mean, and standard deviation of the A's; and n_y, $\bar{x}_y$, and σ_{n_y}, the number of individuals, the mean, and the standard deviation for a y-array of A's. Now clearly, a knowledge of $\bar{y}_x$ and σ_{n_x} will not fix the B's which will be found associated with a given A, but it will define the limits of probable or even possible B's. The curve obtained by plotting $\bar{y}_x$ to x is termed the regression curve of y on x. A curve in which the ratio of σ_{n_x} to the standard deviation σ_y is plotted to x may be termed a scedastic curve*. Since the standard deviation is always a positive quantity, this curve always lies on one side of the axis. It is a horizontal line in the case of normal correlation† and coincides with the axis in any case where correlation passes into causation"‡.

3. Various concepts of correlation held by the English school.

(a) The most important measure of the degree of correlation between two given series or curves x and y, is the Pearsonian (or "product-moment") *coefficient of correlation* r, which may be defined by the simple formula§

$$(1)\quad r = \frac{(1/n)\sum_1^n{}_i (x_i y_i) - \bar{x}\,\bar{y}}{\left\{(1/n)\sum_1^n{}_i (x_i^2) - \bar{x}^2\right\}^{1/2}\left\{(1/n)\sum_1^n{}_i (y_i^2) - \bar{y}^2\right\}^{1/2}} = \frac{(1/n)\sum_1^n{}_i X_i Y_i}{\sigma_x \sigma_y}.$$

It is readily seen from (1) that we may interpret r as the cosine of the angle between two vectors in Euclidian n-space; and thus r is always such that $-1 \leqq r \leqq 1$.

We may also write that

$$(2)\qquad r = \left\{1 - \frac{s_y^2}{\sigma_y^2}\right\}^{1/2}$$

where s_y^2 is the arithmetic mean of the squares of the deviations of the dots of the scatter-diagram from the *line of regression* of y on x which

* A curve which measures the "scatter" in the arrays.

† The Gauss-Laplacian distribution of deviations.

‡ When one value of B only is associated with each A.

§ Whittaker & Robinson, "The Calculus of Observations" (1924), pp. 317–342.

Rietz, H. L., "Mathematical Statistics", Carus Mathematical Monograph No. 3 (1927), pp. 77–113.

Yule, G. U., Proc. Roy. Soc., vol. 60 (1897), p. 477; "Introduction to the Theory of Statistics" (1915), pp. 168–175.

Rietz, H. L., and Others: "Handbook of Mathematical Statistics" (1924), p. 27.

Rietz, H. L., Quart. Publ. Amer. Statistical Assoc. vol. 16 (1919), pp. 472–476.

Huntington, E. V., Amer. Math. Monthly, vol. 26 (1919), pp. 424–434.

Weida, F. M., "On the correlation between two functions", Amer. Math. Monthly, vol. 33 (1926), pp. 440–444.

has the property that the sum of the squares of the distances (measured parallel to the y-axis) of all the dots is less than from any other line, and σ_y^2 is the arithmetic mean of the deviations of the y's from their arithmetic mean.

In the Pearsonian sense, perfect correlation (positive or negative) between two sets of quantities x and y means that there exists a linear algebraic equation connecting those quantities. In general, the given sets of values are not linearly related and the value of r will be less than 1. It is easily shown that the value of r is a suitable criterion of the approach to linear relationship.

(b) The Millsian *indices of correlation* r_{yx} and r_{xy} are defined by*

$$(3) \qquad r_{yx} = \left\{1 - \frac{s_y^2}{\sigma_y^2}\right\}^{1/2} \text{ and } r_{xy} = \left\{1 - \frac{s_x^2}{\sigma_x^2}\right\}^{1/2}$$

where s_y^2 is the arithmetic mean of the squares of the deviations of the dots of the scatter-diagram from the *curve of regression* of y on x which has the property that the sum of the squares of the distances (measured parallel to the y-axis) of all the dots is less than from any other similar curve, s_x^2 is the arithmetic mean of the squares of the deviations of the dots from the *curve of regression* of x on y, and σ_x^2 is the arithmetic mean of the squares of the deviations of the x's from their arithmetic mean.

In the Millsian sense, perfect correlation between two sets of quantities x and y means that there exists a non-linear equation connecting those quantities which is always a single-valued function. In general, the given sets of values are not connected by a non-linear single-valued function and the values r_{yx} and r_{xy} will be less than 1. It is readily seen that r_{yx} and r_{xy} are suitable criterions of the approach to non-linear single-valued relationship.

(c) It was in connection with the general conception of correlation mentioned in article 2 of this paper that Pearson first gave the *correlation ratios* η_{yx} and η_{xy} as more general measures than the correlation coefficient. This conception did not make the correlation of y with x depend solely on the variations in the mean value of the y's in the arrays which correspond to different assigned values of x; but also considered the variations in the standard deviation† of the y's in these arrays. We may make a similar statement for the conception in regard to the correlation of x with y.

---

* Mills, F. C., "Statistical Methods", (1924), pp. 432–441.

† Rietz, H. L., and Others, "Handbook of Mathematical Statistics" (1924), p. 27; also pp. 129–131.

Let s_{ay} denote the standard deviation of any array of y's, and let n_x be the number of observations in this array, and let $\sigma^2_{ay} = \sum (n_x s^2_{ay})/n$. Then*

$$(4) \qquad \eta^2_{xy} = 1 - \frac{\sigma^2_{ay}}{\sigma^2_y}, \text{ and similarly } \eta^2_{xy} = 1 - \frac{\sigma^2_{ax}}{\sigma^2_x}.$$

The correlation ratio may be considered a special case of the correlation index. It is employed when the regression is not linear, and it measures the degree of relationship between two variables in so far as this relationship may be described by a curve passing through the mean of every array. The correlation ratio is also a measure of the concentration of the dots of the scatter-diagram about the regression curve which passes through the means of the arrays.

The above concepts relate to correlation between items or characters that are capable of quantitative measurement. When data are not measurements expressed in cardinal numbers, we may seek a measure of the correlation of items that are (1) merely marks of the orders or ranks of individuals in a series; (2) variables which admit of exact measurement with great labor such as colors, temperaments, shapes, and so on; (3) variables, one of which is measurable while the other is given in alternate categories; (4) variables, one of which is given in multiple categories while the other is given in alternate categories; (5) attributes; and (6) variables not capable of quantitative measurement. The consideration of association or correlation of items that can not be measured, but can be classified, brought about the development of the notions *correlation from ranks, Spearman's foot-rule, tetrachoric correlation* by Pearson, *bi-serial correlation, measures of association* by Yule, and the *measures of contingency by Pearson.*†

---

* Pearson, Karl, "On the general theory of skew correlation and non-linear regression" Drapers' Co. Research Memoirs: Biometric Series II: London, 1905.

Crathorne, A. R., "Calculation of the correlation ratio", Quart. Publ. Amer. Stat. Assoc., vol. 18 (1922), pp. 394-396.

† Pearson, Karl, "On further methods of determining correlation", Drapers' Company Research Memoirs, Biometric Series IV (1907), p. 13.

Spearman, C., "A foot-rule for measuring correlation", Brit. Jour. of Psychology, vol. 2 (1906), p. 89; also vol. 3 (1910), p. 271.

Biometrica, vol. 9 (1913), pp. 22–27.

Phil. Trans. Roy. Soc., A, vol. 194 (1900), p. 257.

Yule, G. U., "Introduction to the Theory of Statistics", (1915), p. 38.

Jour. Roy. Stat. Soc., vol. 75 (1911–12), pp. 579–652.

Pearson, Karl, "On the theories of association", Biometrica, vol. 9 (1913), pp. 159–315.

Pearson, Karl, "On the theory of contingency and its relation to association and normal correlation", Drapers' Company Research Memoirs, Biometric Series I (1904); "On the measurement of the influence of 'Broad categories' on correlation", Biometrica, vol. 9 (1913),

In devising measures for correlation of characters not capable of quantitative measurement, it seems that Pearson has held pretty closely to the thought that such measures should have the property that they will approach the correlation coefficient if the distributions are given by normal frequency curves and surfaces. On the other hand, Yule does not seem to have held closely to any such principle. Neither does it seem that Gini has made his measures of association conform to this principle.

PART II.

Connection and concordance.

4. Gini has divided the problem of association of two statistical variables into the problem of *connection* and the problem of *concordance*. He regards the correlation ratio as a special measure of connection but not a measure of concordance. In fact, the correlation ratio shows the degree of connection or lack of independence, but since its sign is always positive, it is fairly clear that it does not show whether the two characters are concordant or discordant. In other words, it is not shown by the correlation ratio whether the corresponding modalities are more concordant and less discordant than they would be under independence. On the other hand, Gini classifies the correlation coefficient among his special measures of concordance. When the regression is linear, it is clear that a positive correlation coefficient indicates that the characters are, to a certain extent, concordant, and a negative correlation coefficient indicates that they are discordant. In other words, two characters are said to be connected each with the other when the distribution of the modalities of the first one depends on the modalities of the other. The problem of concordance is the determination whether the modalities of the first character are more frequently associated to the concordant modalities, or to the discordant ones of the second character. It is fairly obvious that there is not concordance if connection does not exist.

Before discussing the relations between qualitative and quantitative modalities of two characters, we will present the fundamental principles

---

pp. 116–139 and 216–217; "On a new method of determining correlation etc.", Biometrica, vol. 7 (1909–10), pp. 96–105, 248–257.

Elderton, W. P., "Frequency Curves and Correlation", (1906), pp. 125–131.

Soper, H. E., Biometrica, vol. 10 (1914–15), pp. 384–390.

Rietz, H. L., and Others: "Handbook of Mathematical Statistics", (1924), pp. 132–140.

Pearson, Karl, Proc. Roy. Soc., vol. 60 (1897), p. 489; Phil. Trans., vol. 192 (1899), p. 257.

Yule, G. U., Jour. Roy. Stat. Soc., vol. 73 (1910), p. 644; "Introduction to the Theory of Statistics" (1917), pp. 218–219.

Phil. Trans., A, Vol. 195 (1910), pp. 1–47.

upon which the concepts of connection and concordance are based, namely, *variability and dissimilarity.*

5. **Indices of variability.** Let us consider a set of quantities $x_i (i = 1, 2, \cdots, n)$, and let $\bar{x} = (1/n) \sum_1^n{}_i x_i$; then, the *simple mean* or *average deviation* of the x's is defined as

$$D_x = \frac{\sum_1^n{}_i |x_i - \bar{x}|}{n}, \tag{5}$$

and the *quadratic mean* or *standard deviation* of the x's is defined as

$$\sigma_x = \left\{ \frac{\sum_1^n{}_i (x_i - \bar{x})^2}{n} \right\}^{1/2}. \tag{6}$$

If we have given a series of n ordered quantities x_i, such that $x_1 \leqq x_2 \leqq x_3 \leqq \cdots \leqq x_n$; then we shall say that the arithmetic mean of the absolute values of all the differences which may be calculated among the terms of the series is the *mean difference.* If, in this calculation, we compare a term with itself, then we shall have the *mean difference with repetition*; if we do not compare a term with itself, then we shall have the *simple mean difference.*

Let us now form the sums

$$\begin{aligned}
s_1 &= x_1, & s_1' &= x_n, \\
s_2 &= x_1 + x_2, & s_2' &= x_n + x_{n-1}, \\
&\cdots\cdots\cdots & &\cdots\cdots\cdots \\
s_n &= x_1 + x_2 + \cdots + x_n, & s_n' &= x_n + x_{n-1} + \cdots + x_1, \\
S &= s_1 + s_2 + \cdots + s_n, & S' &= s_1' + s_2' + \cdots + s_n',
\end{aligned}$$

then, the *simple mean difference* $({}_mD_x)$—the mean difference between the x's without repetition—is given by

$$({}_mD_x) = \frac{S' - S}{n(n-1)/2}. \tag{7}$$

The *t-ic mean difference* ${}^t({}_mD_x)_R$ between the x's *with repetition* is defined by

$${}^t({}_mD_x)_R = \left\{ \frac{\sum_{ij} |x_i - x_j|^t}{n^t} \right\}^{1/t}. \tag{8}$$

If $t = 1$, we obtain from (8) the *mean difference* between x's *with repetition,* namely,

$$(9) \qquad {}_{(m}D_{x)R} = \frac{2(S'-S)}{n^2} = \frac{n-1}{n}({}_{m}D_{x}),$$

and if $t = 2$, we obtain from (8) the *quadratic mean difference* between the x's *with repetition*, namely,

$$(10) \qquad {}^{2}({}_{m}D_{x})_{R} = \left\{\frac{\sum_{ij}|x_i - x_j|}{n^2}\right\}^{1/2} = \sigma_x(2)^{1/2}.$$

6. **Dissimilarity.** Let A and B be two groups of intensities of a quantity x which is defined for a set of values $x = x_i\,(i = 1, 2, \cdots, k)$. Let f_i be the frequency of x_i in the group A of m quantities and let g_i be the frequency of x_i in the group B of n quantities.

The two groups A and B are said to be *similar* if $f_i/g_i = m/n$ for all values of i. If A and B are similar and if $m = n$, then $f_i = g_i$ for every i; and obviously, all the quantities of A may be paired to the quantities of B so that the difference between paired quantities will disappear. Again, if A and B are similar or dissimilar and composed of a different number of quantities, it is fairly clear that we can always form a group A' similar to A and a group B' similar to B, both A' and B' consisting of say n' quantities. Hence, when we discuss the similarity of two groups, we need only consider groups composed of the same number of quantities.

To obtain a method for the comparison of the quantities of two groups, we define (1) the concept of *cograduate intensities* and (2) the concept of *contragraduate intensities.*

Definition (1). Two intensities are said to be *cograduate* when they present the same degree (the ordinal number which the intensity has in the graduation) in two graduations in which both are increasing or decreasing.

Definition (2). Two intensities are said to be *contragraduate* when they present the same degree in two graduations, one of which is increasing and the other is decreasing.

Thus, considering the intensities of a character according to an *increasing* or *decreasing* order of graduation, if

$$(11) \qquad x_1 \geqq x_2 \geqq \cdots \geqq x_r \geqq \cdots \geqq x_{n-r+1} \geqq \cdots \geqq x_n$$

is the A group of ordered intensities of the quantity x and if

$$(12) \qquad y_1 \geqq y_2 \geqq \cdots \geqq y_r \geqq \cdots \geqq y_{n-r+1} \geqq \cdots \geqq y_n$$

is the B group of ordered intensities of the quantity x, then the quantities x_r and y_r are said to be cograduate intensities and the quantities x_r and y_{n-r+1} or x_{n-r+1} and y_r are said to be contragraduate intensities.

It is fairly easy to show that if, in the two groups (11) and (12) of ordered quantities, we pair x_r with y_r, then the sum $\sum_1^n{}_r |x_p - y_p|$ will be a minimum: and if we pair x_r with y_{n-r+1}, then the sum $\sum_1^n{}_r |x_r - y_{n-r+1}|$ will be a maximum. Likewise, it can be shown that the sum $\sum_1^n{}_r (x_r - y_r)^2$ will be a minimum and the sum $\sum_1^n{}_r (x_r - y_{n-r+1})^2$ will be a maximum.

To obtain a measure of dissimilarity, it appears natural to compare the quantities according to a correspondence which makes the absolute value of the sum of the differences between compared quantities a minimum.* Accordingly, we take the *arithmetical mean* of these differences as the *simple index of dissimilarity,* and the *quadratic mean* as the *quadratic index of dissimilarity.*

Hence, the *simple index of dissimilarity* d is defined by

$$d = (1/n) \sum_1^n{}_i |x_i - y_i|, \tag{13}$$

and the *quadratic index of dissimilarity* ${}^2d$

$${}^2d = \left\{(1/n) \sum_1^n{}_i (x_i - y_i)^2\right\}^{1/2}. \tag{14}$$

Let $\bar{x} = (1/n) \sum_1^n{}_i x_i$ be the arithmetic mean of the n intensities of a character A, then the value $\Delta_i = x_i - \bar{x}$ is the positive or negative deviation of the intensity x_i from the arithmetic mean. If $D_x = (1/n) \sum_1^n{}_i |x_i - \bar{x}|$ $= (1/n) \sum_1^n{}_i |\Delta_i|$, then the value $v_i = \Delta_i / D_x$ is said to be the *variation* in the ith case.†

It is fairly easy to see that the *simple index of dissimilarity between two groups of deviations* is given by the formula

$$\delta = (1/n) \sum_1^n{}_i |\Delta_i - \Delta'_i|, \tag{15}$$

where Δ'_i refers to a character B; and the *quadratic index of dissimilarity between two groups of deviations* is given by the formula

$${}^2\delta = \left\{(1/n) \sum_1^n{}_i (\Delta_i - \Delta'_i)^2\right\}^{1/2}. \tag{16}$$

* Boas, F., "The measurement of differences between variable quantities", Jour. Amer. Stat. Assoc., vol. 18 (1922), no. 140.1.

† In place of the arithmetic mean we may use the quadratic mean or the median.

Likewise, the *simple index of dissimilarity between two groups of variations* will be given by the formula

$$\delta' = (1/n)\sum_1^n{}_i |v_i - v'_i|, \tag{17}$$

and the *quadratic index of dissimilarity between two groups of variations* will be given by the formula

$${}^2\delta' = \left\{(1/n)\sum_1^n{}_i (v_i - v'_i)^2\right\}^{1/2}. \tag{18}$$

The concept of dissimilarity and the indices of dissimilarity given above should be applied only to the relations between series of quantitative characters or to the relations between series whose modalities correspond to a succession of numbers. Hence, the same method is not usable for cyclical series. Pietra* has stated a theory of the dissimilarity between cyclical series.

Let us now establish a measure of connection and also some indices of connection which are applicable to rectilinear and disconnected series.

7. **A measure of connection.** Let n_i $(i = 1, 2, \cdots, n)$ represent the frequency of the modality y_i of the character B and let n'_j $(j = 1, 2, \cdots, m)$ represent the frequency of the modality x_j of the character A. Let n_{ij} represent the frequency with which the ith modality y_i of the character B corresponds to the jth modality x_j of the character A.

The frequencies of B and A will then be distributed according to the following table.

		$x \longrightarrow$					
		x_1	x_2	$\cdots$	x_j	$\cdots$	x_m	Totals
	y_n	$n_{n,1}$	$n_{n,2}$	$\cdots$	$n_{n,j}$	$\cdots$	$n_{n,m}$	n_n
↑	y_{n-1}	$n_{n-1,1}$	$n_{n-1,2}$	$\cdots$	$n_{n-1,j}$	$\cdots$	$n_{n-1,m}$	n_{n-1}
	.	.	.	$\cdots$	.	$\cdots$	.	.
	y_i	$n_{i,1}$	$n_{i,2}$	$\cdots$	$n_{i,j}$	$\cdots$	$n_{i,m}$	n_i
y	.	.	.	$\cdots$	.	$\cdots$	.	.
	y_1	$n_{1,1}$	$n_{1,2}$	$\cdots$	$n_{1,j}$	$\cdots$	$n_{1,m}$	n_1
	Totals	n'_1	n'_2	$\cdots$	n'_j	$\cdots$	n'_m	N

Such a table is known as a correlation or double-entry table and it contains a system of columns and rows each of which is a frequency distribution. The numbers in a column corresponding to an assigned $x = x_j$

* Pietra, G., Metron, vol. 4 (1924-25), pp. 450–558.

form what is often called an x-array of y's, and those in a row corresponding to $y = y_i$ a y-array of x's. Any number n_{ij} in a compartment of this table is called a cell frequency.

From the above table we see that the frequencies of B are distributed according to m "partial" groups which correspond to the modalities of A.

If all the "partial" groups

$$\begin{aligned} &n_{n1},\ n_{n-1,1},\ \cdots,\ n_{i1},\ \cdots,\ n_{11},\\ &n_{n2},\ n_{n-1,2},\ \cdots,\ n_{i2},\ \cdots,\ n_{12},\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ &n_{nm},\ n_{n-1,m},\ \cdots,\ n_{im},\ \cdots,\ n_{1m} \end{aligned}$$

are similar to the "total" group

$$n_n,\ n_{n-1},\ \cdots,\ n_i,\ \cdots,\ n_1$$

of frequencies of B, then the distribution of the modalities of B is independent of the modalities of A, and B is not connected with A. Again, if at least one of the above "partial" groups is not similar to the "total" group of frequencies of B, then the distribution of the modalities of B is dependent on the modalities of A, and B is connected with A. In other words, the greater the d_j indices of dissimilarity between the "partial" groups and the "total" group of frequencies of B, the greater will be the connection of B on A.

A weighted arithmetic mean of these indices of dissimilarity may be taken as a measure of connection. If I_{yx} represents the measure of connection of B on A, we may write that

$$I_{yx} = (1/N)\sum_{1}^{m}{}_j\, n'_j\, d_j. \tag{19}$$

The measure of connection I_{yx} reaches its *absolute maximum* value $({}_mD_y)_R$ — the mean difference with repetition between the y's — when the total number of intensities of B is equal or inferior to that of the modalities of A. In all other cases, the maximum value of $I_{yx} = M_{yx} < ({}_mD_y)_R$. This latter value of I_{yx} is known as the *relative maximum* and it may be computed by grouping the frequencies of y in such a way that each group contains modalities of x which differ from each other the least amount possible. To illustrate: Considering the above table of frequencies, we first find the arithmetic mean $\overline{y}_x$ for each of the original m "partial" groups of frequencies of B which correspond to the modalities x of A; secondly, we take these found values of the $\overline{y}_x$ and consider them as a set and arrange them in either an ascending or descending order of numerical values; thirdly, we distribute the frequency of each modality y

of B amongst the modalities x of A so that the arithmetic means $\bar{y}_x$ of the partial groups corresponding to the modalities x of A which receive some of the frequencies of the modalities y of B differ from each other the least amount possible in the same direction and so that both the frequencies of the modalities x of A and y of B remain unchanged; and fourthly, from this new table of distributions of the frequencies of B and A, we obtain its measure of connection of B on A, which will give us the desired *relative maximum* M_{yx} of the connection of B on A. Hence, if d'_j, $(j = 1, 2, \cdots, m)$, represents the set of indices of dissimilarity between the new m "partial" groups and the "total" group, then the *relative maximum* of the connection B on A is given by

$$M_{yx} = (1/N) \sum_1^m{}_j \, n'_j \, d'_j \; < ({}_mD_y)_R. \tag{20}$$

Similarly, the measure of connection of A on B is given by

$$I_{xy} = (1/N) \sum_1^n{}_i \, n_i \, d_i. \tag{21}$$

Here, the *absolute maximum* of the connection is $({}_mD_x)_R$ and the *relative maximum* M_{xy} is given by

$$M_{xy} = (1/N) \sum_1^n{}_i \, n_i \, d'_i \; < ({}_mD_x)_R. \tag{22}$$

It is now our purpose to establish certain indices of connection between two characters A and B.

8. **Indices of connection.** To obtain an index of connection between two characters A and B it appears natural to form the ratio of the measure of connection to the maximum value which the connection may attain. Hence, if we take the *absolute maximum* as the maximum, then the *simple indices of connection* C_{yx} of B on A and C_{xy} of A on B are given by

$$C_{yx} = \frac{\sum_1^m{}_j \, n'_j \, d_j}{N({}_mD_y)_R}; \qquad C_{xy} = \frac{\sum_1^n{}_i \, n_i \, d_i}{N({}_mD_x)_R}; \tag{23}$$

and if we take the *relative maximum* for the maximum, then the *simple indices of connection* C'_{yx} of B on A and C'_{xy} of A on B are given by

$$C'_{yx} = \frac{\sum_1^m{}_j \, n'_j \, d_j}{NM_{yx}}; \qquad C'_{xy} = \frac{\sum_1^n{}_i \, n_i \, d_i}{NM_{xy}}. \tag{24}$$

Again, if we take the *absolute maximum* as the maximum, then the *quadratic indices of connection* ${}^2C_{yx}$ of B on A and ${}^2C_{xy}$ of A on B are given by

$$(25) \qquad ({}^2C_{yx})^2 = \frac{\sum_1^m {}_j n_j' ({}^2d_j)^2}{2N\sigma_y^2}; \qquad ({}^2C_{xy})^2 = \frac{\sum_1^n {}_i n_i ({}^2d_i)^2}{2N\sigma_x^2};$$

and if we take the *relative maximum* for the maximum, then the *quadratic indices of connection* ${}^2C'_{yx}$ of B on A and ${}^2C'_{xy}$ of A on B are given by

$$(26) \qquad ({}^2C'_{yx})^2 = \frac{\sum_1^m {}_j n_j' ({}^2d_j)^2}{N({}^2M_{yx})^2}; \qquad ({}^2C'_{xy})^2 = \frac{\sum_1^n {}_i n_i ({}^2d_i)^2}{N({}^2M_{xy})^2};$$

where $({}^2M_{yx}) = \left\{(1/N)\sum_1^m {}_j n_j' ({}^2d_j')^2\right\}^{1/2}$, and $({}^2M_{xy}) = \left\{(1/N)\sum_1^n {}_i n_i ({}^2d_i')^2\right\}^{1/2}$.

The question in regard to the choice of the *relative maximum* rather than the *absolute maximum* depends on the nature of the investigation and is discussed by Gini in "Di una misura della dissomiglianza tra due gruppi di quantita e della sua applicazione allo studio delle relazioni statistiche".

We shall now show that Pearson's *correlation ratio* η is a measure of connection. To introduce a convenient notation, let $\bar{y}$ represent the arithmetic mean of the total group of n intensities of y, and let $\bar{y}_{x_j}$ $(j = 1, 2, \cdots, m)$ represent the arithmetic mean of the partial group of n_j' intensities of y corresponding to the jth modality of x; then the mean μ_{yx} of the absolute value of the difference between the means of the x-array of y's and the mean of the total group of y intensities is given by

$$(27) \qquad \mu_{yx} = (1/N)\sum_1^m {}_j n_j' |\bar{y}_{x_j} - \bar{y}|.$$

The *absolute maximum* of μ_{yx} is the simple mean deviation D_y. This maximum is attained only in the case when to each modality x of A there always corresponds the same modality y of B. Similarly, the mean μ_{xy} of the absolute value of the difference between the means of the y-array of x's and the mean of the total group of x intensities is given by

$$(28) \qquad \mu_{xy} = (1/N)\sum_1^n {}_i n_i |\bar{x}_{y_i} - \bar{x}|,$$

and the absolute maximum of μ_{xy} is D_x.

Hence, the *simple indices of connection* $({}_mC_{yx})$ *between means* of B on A and $({}_mC_{xy})$ *between means* of A on B referred to the *absolute maximum* are given by

$$({}_mC_{yx}) = (1/ND_y)\sum_1^m{}_j n'_j|\bar{y}_{x_j}-\bar{y}|, \tag{29}$$

$$({}_mC_{xy}) = (1/ND_x)\sum_1^n{}_i n_i|\bar{x}_{y_i}-\bar{x}|. \tag{30}$$

The *quadratic indices of connection* ${}^2({}_mC_{yx})$ *between means* of B on A and ${}^2({}_mC_{xy})$ *between means* of A on B referred to the absolute maximum are given by

$$\{{}^2({}_mC_{yx})\}^2 = \frac{\sum_1^m{}_j n'_j(\bar{y}_{x_j}-\bar{y})^2}{N\sigma_y^2} = \frac{\sigma_{my}^2}{\sigma_y^2}, \tag{31}$$

$$\{{}^2({}_mC_{xy})\}^2 = \frac{\sum_1^n{}_i n_i(\bar{x}_{y_i}-\bar{x})^2}{N\sigma_x^2} = \frac{\sigma_{mx}^2}{\sigma_x^2}; \tag{32}$$

where σ_{my} is the standard deviation of the $\bar{y}_{x_j}$'s and σ_{mx} is the standard deviation of the $\bar{x}_{y_i}$'s.

The expressions (31) and (32) are Pearson's *correlation ratios.* To establish this fact we apply the well known theorem: "In any single array, the sum of the squares of the deviations from any axis, say $\bar{y}$, is equal to the sum of the squares of the deviations from a parallel axis, say $\bar{y}_{x_j}$, passing through the mean of the array, plus the product of the number in the array and the square of the distance between the two axes, namely, the deviation $\bar{y}_{x_j}-\bar{y}$."

For all the arrays, we find that

$$N\sigma_y^2 = \sum_1^m{}_j n'_j(\sigma_{x_j}^2+[\bar{y}_{x_j}-\bar{y}]^2). \tag{33}$$

Rewriting (33) we find that

$$\sigma_{my}^2 = \sigma_y^2-\sigma_{ay}^2. \tag{34}$$

From (4) and (34) we find that

$$\eta_{yx}^2 = \frac{\sigma_{my}^2}{\sigma_y^2}. \tag{35}$$

From (35) and (31) we see that

$$\eta_{yx}^2 = \{{}^2({}_mC_{yx})\}^2. \tag{36}$$

Similarly, it can be shown that

$$\eta^2_{xy} = \{^2({}_mC_{xy})\}^2. \tag{37}$$

Hence, we see that Pearson's *correlation ratios* are indices of a special case of connection between two groups.

We will now state certain principles, established by Gini, for determining whether the modalities x of A are more frequently associated to the concordant modalities or to the discordant modalities y of B.

9. **Principles of concordance.** We have the following principles:

(a) Modalities (deviations, variations) of two characters may be called concordant or discordant, when the discordances between corresponding modalities (deviations, variations) are respectively lower or higher than those which would result if the modalities (deviations, variations) of a character would have been associated by chance to those of the other one.

(b) Modalities (deviations, variations) of two characters may be called concordant or discordant, when the discordances between corresponding modalities (deviations, variations) are respectively lower or higher than those between modalities (deviations, variations) of a character and the opposite of the other one.

(c) Modalities (deviations, variations) of two characters may be called concordant or discordant, when, subtracting from the sum of the discordances between the corresponding modalities (deviations, variations) of two characters, its probable value in the hypothesis of associations by chance, the result will be lower or higher than that which may be obtained by subtracting the probable value, in the hypothesis of associations by chance, from the sum of the discordances between the corresponding modalities (deviations, variations) of a character and the opposite of the other one.

(d) Modalities (deviations, variations) of two characters A and B will be concordant, or discordant, when, subtracting from the sum of the discordances between corresponding modalities (deviations variations) of A and B, and of discordances between corresponding modalities (deviations, variations) of a and b, (a being the opposite of A, and b the opposite of B), its probable value in the hypothesis of association by chance, the result will be lower or higher than that which may be obtained by subtracting the probable value in the hypothesis of association by chance, from the sum of the discordances between corresponding modalities (deviations, variations) of a and B and of the discordances between corresponding modalities (deviations, variations) of A and b.

We shall now form analytic formulas for the measure of concordance between modalities (deviations, variations) of two characters A and B according to each of the above principles.

10. **The measure of concordance.** To introduce a convenient notation, let M_{AB}, D_{AB}, and V_{AB} represent respectively the sum of the discordances between corresponding modalities (deviations, variations) of two characters A and B. Let M'_{AB}, D'_{AB}, and V'_{AB} represent the respective probable values in the hypothesis of association by chance. Let analogous symbols represent the above sums in the case of the opposite characters a and b. Thus:

According to principle (a) *of concordance*: the modalities (deviations, variations) of A and B are concordant if $M_{AB} - M'_{AB} < 0$, $D_{AB} - D'_{AB} < 0$, $V_{AB} - V'_{AB} < 0$; and discordant if $M_{AB} - M'_{AB} > 0$, $D_{AB} - D'_{AB} > 0$, $V_{AB} - V'_{AB} > 0$.

According to principle (b) *of concordance*: the modalities (deviations, variations) of A and B are concordant if $M_{AB} - M_{Ab} < 0$, $D_{AB} - D_{Ab} < 0$, $V_{AB} - V_{Ab} < 0$; and discordant if $M_{AB} - M_{Ab} > 0$, $D_{AB} - D_{Ab} > 0$, $V_{AB} - V_{Ab} > 0$.

According to principle (c) *of concordance*: the modalities (deviations, variations) of A and B are concordant if $(M_{AB} - M'_{AB}) - (M_{Ab} - M'_{Ab}) < 0$, $(D_{AB} - D'_{AB}) - (D_{Ab} - D'_{Ab}) < 0$, $(V_{AB} - V'_{AB}) - (V_{Ab} - V'_{Ab}) < 0$; and discordant if $(M_{AB} - M'_{AB}) - (M_{Ab} - M'_{Ab}) > 0$, $(D_{AB} - D'_{AB}) - (D_{Ab} - D'_{Ab}) > 0$, $(V_{AB} - V'_{AB}) - (V_{Ab} - V'_{Ab}) > 0$.

According to principle (d) *of concordance*: the modalities (deviations, variations of A and B are concordant if $(M_{AB} + M_{ab} - M'_{AB} - M'_{ab}) - (M_{Ab} + M_{aB} - M'_{Ab} - M'_{aB}) < 0$, $(D_{AB} + D_{ab} - D'_{AB} - D'_{ab}) - (D_{Ab} + D_{aB} - D'_{Ab} - D'_{aB}) < 0$, $(V_{AB} + V_{ab} - V'_{AB} - V'_{ab}) - (V_{Ab} + V_{aB} - V'_{Ab} - V'_{aB}) < 0$; and discordant if $(M_{AB} + M_{ab} - M'_{AB} - M'_{ab}) - (M_{Ab} + M_{aB} - M'_{Ab} - M'_{aB}) > 0$, $(D_{AB} + D_{ab} - D'_{AB} - D'_{ab}) - (D_{Ab} + D_{aB} - D'_{Ab} - D'_{aB}) > 0$, $(V_{AB} + V_{ab} - V'_{AB} - V'_{ab}) - (V_{Ab} + V_{aB} - V'_{Ab} - V'_{aB}) > 0$.

The above inequalities represent measures of concordance. If these measures define a negative quantity we say that the characters A and B are concordant and if they define a positive quantity we say that the characters A and B are discordant. Hence, the greater the concordance the greater will be the negative quantities and the greater the discordance the greater will be the positive quantities.

We have stated four principles of concordance. In fact, there is but one principle of concordance and it is principle (d) of concordance given above. By comparing the four principles of concordance given above, we readily see that principles (a), (b), and (c) are special cases of principle (d). The availability for use of each of the principles (a), (b), and (c) is dependent

upon the nature of the statistical investigation while principle (d) is always available.

The unit of measure may be an *absolute maximum* or a *relative maximum*. The *relative maximum* is such that it may be verified compatibly with the effective distribution of the intensities of two characters A and B. The *absolute maximum* is the maximum of the relative maxima.

The maximum intensity that an index of concordance may reach is $+1$, and the maximum intensity that an index of discordance may reach is -1. In the case of independence we have zero.

If we use the *relative maximum* and *principle* (a) *of concordance* when we are dealing with characters capable of quantitative measure we obtain Gini's *indices of homophilia*. If we use the *absolute maximum* and *principle* (a) *of concordance* when we are dealing with characters capable of quantitative measure we obtain Gini's *indices of correlation*. If we use the *relative maximum* and *principle* (a) *of concordance* when we are dealing with qualitative characters we obtain Gini's *indices of attraction*. If we use the *absolute maximum* and *principle* (a) *of concordance* when we are dealing with *qualitative characters* we obtain Gini's *indices of likeness*. We may obtain both *simple* and *quadratic* indices. In the case of *simple indices* we measure the unlikeness of two quantities by the absolute value of their difference and in the case of *quadratic indices* we measure the unlikeness of two quantities by the square of the differences.

To illustrate the concept of concordance we will now show how Gini's indices of homophilia, correlation, attraction, and likeness may be obtained.

11. **Indices of homophilia.** If ${}^{c}_{r}M_{BA}$ is the value of M_{BA} in the hypothesis of relative maximum of concordance between the modalities of two characters B and A, then ${}^{c}_{r}M_{BA}$ is the minimum of the sum of the absolute values of the differences between corresponding intensities of the characters B and A. Again, if ${}^{d}_{r}M_{BA}$ is the value of M_{BA} in the hypothesis of relative maximum of discordance between the modalities of two characters B and A, then ${}^{d}_{r}M_{BA}$ is the maximum of the sum of the absolute values of the differences between corresponding intensities of the two characters B and A.

Hence, the *simple index of homophilia*, according to principle (a) of concordance, between the modalities of B and A is

$$H_M = \frac{M_{BA} - M'_{BA}}{{}^{c}_{r}M_{BA} - M'_{BA}} \text{ or } - \frac{M_{BA} - M'_{BA}}{{}^{d}_{r}M_{BA} - M'_{BA}} \tag{38}$$

according as $M_{BA} \gtrless M'_{BA}$.

The probable value M'_{BA} of M_{BA} is given by the formula

$$M'_{BA} = N({}_{m}D_{yx})_R = N\{2({}_{m}D_{\overline{y+x}})_R - 1/2)[({}_{m}D_x)_R + ({}_{m}D_y)_R]\}, \tag{39}$$

where $({}_mD_x)_R$, $({}_mD_y)_R$, $({}_mD_{\overline{y+x}})_R$ are respectively the mean differences with repetition between the N intensities of A, the N intensities of B, and the $2N$ intensities of B and A together.

Referring to the double-entry table given in article 7 of this paper, it is easy to see that the sum of the discordances M_{BA}, between corresponding modalities of B and A is given by

$$M_{BA} = \sum_{ij} n_{ij} \left| x_j - y_i \right|. \tag{40}$$

The values of ${}^c_rM_{BA}$ and ${}^d_rM_{BA}$ are given by

$${}^c_rM_{BA} = \sum_r \left| x_r - y_r \right| \text{ and } {}^d_rM_{BA} = \sum_r \left| x_r - y_{n-r+1} \right|, \tag{41}$$

where x_r and y_r are cograduate intensities of A und B, and x_r and y_{n-r+1} are contragraduate intensities of A and B.

If we now substitute the values given by (39), (40) and (41) for the quantities indicated in (38) we find that H_M is given by

$$H_M = \frac{\sum_{ij} n_{ij} \left| x_j - y_i \right| - N({}_mD_{yx})_R}{\sum_1^n{}_r \left| x_r - y_r \right| - N({}_mD_{yx})_R}, \tag{42}$$

or

$$H_M = \frac{N({}_mD_{yx})_R - \sum_{ij} n_{ij} \left| x_j - y_i \right|}{\sum_1^n{}_r \left| x_r - y_{n-r+1} \right| - N({}_mD_{yx})_R}, \tag{43}$$

according as $\sum_{ij} n_{ij} \left| x_j - y_i \right| \leqq N({}_mD_{yx})_R$.

Similarly, formulas for the *simple index of homophilia* between the modalities of two characters B and A according to principles (b), (c), and (d) of concordance may be established. Formulas, analogous to (42) and (43), may also be obtained to define the *simple indices of homophilia* between deviations and variations of two characters B and A.

If ${}^2H_M$ represents the *quadratic index of homophilia* between modalities, if ${}^2H_D$ represents the *quadratic index of homophilia* between deviations, and if ${}^2H_V$ represents the *quadratic index of homophilia* between variations; then, assuming principle (a) of concordance and measuring the unlikeness of the quantities by the square of their differences, we find that

$${}^2H_M = {}^2H_D = {}^2H_V = \frac{\sum_{ij} n_{ij} (x_j - \overline{x})(y_i - \overline{y})}{\sum_1^n{}_r (x_r - \overline{x})(y_r - \overline{y})} = \frac{\sum_1^n{}_k x_k y_k - N\overline{x}\,\overline{y}}{\sum_1^n{}_r \Delta_r \Delta'_r}, \tag{44}$$

or

$$(45)\quad {}^2H_M = {}^2H_D = {}^2H_V = \frac{-\sum_{ij} n_{ij}(x_j-\bar{x})(y_i-\bar{y})}{\sum_1^n{}_r (x_r-\bar{x})(y_{n-r+1}-\bar{y})} = \frac{N\bar{x}\,\bar{y}-\sum_1^{n_s}{}_k x_k y_k}{\sum_1^n{}_r \Delta_r \Delta'_{n-r+1}},$$

where x_k and y_k are pairs of corresponding intensities and x_r and y_r are cograduate intensities and x_r and y_{n-r+1} are contragraduate intensities.

We may always assume principle (a) of concordance to obtain the *quadratic indices of homophilia* because formulas (44) and (45) are always capable of being used for the purpose intended.

We recall that in obtaining the above *indices of homophilia* we made use of the relative maximum. We will now assume principle (a) of concordance and will use the absolute maximum to obtain the so-called *indices of correlation.*

12. **Indices of correlation.** If we refer the measure of concordance to the absolute maximum it can attain and if we use the absolute values of the differences between two quantities to measure their unlikeness; then, according to principle (a) of concordance, the *simple index of correlation* R_M between the modalities of two characters A and B is found to be

$$(46)\quad R_M = \frac{N({}_mD_{yx})_R - \sum_{ij} n_{ij}|x_j-y_i|}{N({}_mD_{y+x+x'+y'})_R},$$

or

$$(47)\quad R_M = \frac{N({}_mD_{yx})_R - \sum_{ij} n_{ij}|x_j-y_i|}{2N(D_{y+x+y'+x'}) - N({}_mD_{y+x+y'+x'})_R},$$

according as $N({}_mD_{yx})_R \gtrless \sum_{ij} n_{ij}|x_j-y_i|$ and where y' and x' are respectively the intensities of the characters b the opposite of B and a the opposite of A.

Similarly, formulas for the simple index of correlation between modalities of two characters according to principles (b), (c), and (d) of concordance may be established. Formulas analogous to (46) and (47) may also be obtained to define the *simple indices of correlation* between deviations and variations.

If we refer the measure of concordance to the absolute maximum it can attain and if we use the squares of the differences between two quantities to measure their unlikeness; then according to principle (a) of concordance, the *quadratic index of correlation* ${}^2R_M$ between the modalities of two characters B and A is found to be

$$(48)\quad {}^2R_M = \frac{\sum_{ij} n_{ij}(x_j-\bar{x})(y_i-\bar{y})}{N\left\{\frac{\sigma_x^2+\sigma_y^2}{2} + D^2_{m|x,y,x',y'}\right\}},$$

where $D^2_{m|x,y,x',y'}$ represents the mean of the sum of the squares of the differences between the mean of the A, B, a, and b groups of intensities taken together and the mean of each of the A, B, a, and b groups of intensities.

If we refer the measure of concordance to the absolute maximum it can attain and if we use the squares of the differences between two quantities to measure their unlikeness; then, according to principle (a) of concordance, the *quadratic index of correlation* ${}^2R_D$ between the deviations of two characters B and A is found to be

$$(49) \qquad {}^2R_D = \frac{\sum_{ij} n_{ij}(x_j - \bar{x})(y_i - \bar{y})}{N\left\{\frac{\sigma_x^2 + \sigma_y^2}{2}\right\}}.$$

If we refer the measure of concordance to the absolute maximum it can attain and if we use the squares of the differences between two quantities to measure their unlikeness; then, according to principle (a) of concordance, the *quadratic index of correlation* ${}^2R_V$ between the variations of two characters B and A is found to be

$$(50) \qquad {}^2R_V = \frac{\sum_{ij} n_{ij}(x_j - \bar{x})(y_i - \bar{y})}{N\sigma_x \sigma_y}.$$

It is easy to see, if we refer to the table given in article 7 of this paper, that $\sum_{ij} n_{ij}(x_j - \bar{x})(y_i - \bar{y}) = \sum_1^n{}_k x_k y_k - N\bar{x}\bar{y}$ where x_k and y_k are pairs of corresponding intensities.

Hence, we may write for (50) that

$$(51) \quad {}^2R_V = \frac{(1/N)\sum_1^n{}_k x_k y_k - \bar{x}\bar{y}}{\sigma_x \sigma_y} = \frac{(1/N)\sum_1^n{}_k X_k Y_k}{\sigma_x \sigma_y} = (1/N)\sum_1^n{}_k X'_k Y'_k.$$

If we compare formula (51) with formula (1), we see at once that (51) is the formula for the well known Pearsonian coefficient of correlation. Hence, the *correlation coefficient* r may be considered a *measure of concordance.*

We may always assume principle (a) of concordance to establish the quadratic indices of correlation because formulas (48), (49), (50), and (51) are always capable of being used for the purpose intended.

We will now study the concordance between two qualitative characters.

13. **Indices of likeness.** To introduce a convenient notation, let s represent the number of modalities of a character A, and also of a character B.

Let n'_i represent the number of cases in which the ith modality appears in A, let n_i represent the number of cases in which the ith modality appears in B, and let n_{ii} represent the number of cases in which the ith modality appears simultaneously in A and B. If n represents the total number of observed cases, then

$$n = \sum_1^s{}_i n'_i = \sum_1^s{}_i n_i.$$

Also, if m'_i represents the number of cases of all the modalities of A other than the ith modality and if m_i represents the number of cases of all the modalities of B other than the ith modality; then

$$m'_i = n - n'_i \quad \text{and} \quad m_i = n - n_i.$$

If we use the absolute maximum of concordance, then according to principle (a) of concordance, the *quadratic index of likeness* ${}^2L_M$ between the modalities of two qualitative characters B and A is found to be

$$(52) \qquad {}^2L_M = \frac{-4\sum_1^s{}_i\{(1/n)n'_i n_i - n_{ii}\}}{sn}.$$

Similarly, according to principle (a) of concordance, the quadratic index of likeness ${}^2L_D$ between the deviations of two characters B and A is found to be

$$(53) \qquad {}^2L_D = \frac{2\sum_1^s{}_i\{(1/n)\, n'_i\, n_i - n_{ii}\}}{2n - (1/n)\sum_1^s{}_i (n'_i)^2 - (1/n)\sum_1^s{}_i (n_i)^2},$$

and the quadratic index of likeness ${}^2L_{\overline{V}}$ between the variations of two characters B and A is found to be

$$(54) \qquad {}^2L_V = \frac{\sum_1^s{}_i\{(1/n)\, n'_i\, n_i - n_{ii}\}}{\left[\left\{n - (1/n)\sum_1^s{}_i (n'_i)^2\right\} \times \left\{n - (1/n)\sum_1^s{}_i (n_i)^2\right\}\right]^{1/2}}.$$

The indices of likeness for qualitative characters are analogous to the indices of correlation for quantitative characters. Examples of calculating ${}^2L_M$, ${}^2L_D$, and ${}^2L_V$ may be found in a paper by Prof. Savorgnan, namely, "La scelta matrimoniale", Biblioteca del Metron, serie B, $n^r - 2$, Taddei, Ferrari.

14. **Indices of attraction.** Using the notation of article 13, the quadratic indices of attraction ${}^2A_M$, ${}^2A_D$, and ${}^2A_V$ between modalities, deviations, and variations of two characters B and A are found to be given, according to a negative or positive numerator, by

$$(55) \qquad {}^2A = \frac{\sum_1^s{}_i \{(1/n)\, n_i' n_i - n_{ii}\}}{\sum_1^s{}_i \{(1/n)\, n_i' n_i - (1/2)\, (n_i' + n_i + |n_i' - n_i|/2)\}},$$

or

$$(56) \qquad {}^2A = \frac{\sum_1^s{}_i \{(1/n)\, n_i' n_i - n_{ii}\}}{\sum_1^s{}_i \{(1/n)\, (n_i' + m_i - |n_i' - m_i|) - (1/n)\, n_i' m_i\}}.$$

The indices of attraction for qualitative characters are analogous to the indices of homophilia for quantitative characters.

15. **Remarks on further concepts of correlation belonging to the English and Italian schools.** At this point, it may be of interest to call attention to certain omissions in the discussion given in Parts I and II, and to give references for further reading.

We have merely mentioned, with references, the concepts of correlation of qualitative characters—characters which do not seem to admit of exact measurement, but admit of classification. We have made no statement concerning the methods of dealing with the concept of correlation* in time series—series of aggregates, or averages, or relative numbers applying to a definite interval or point of time. No discussion has been given of the theory of multiple† correlation—correlation between three or more variables; and we have not mentioned the theory of partial‡ correlation—correlation between two variables x_1 and x_2 when the other variables $x_3, x_4, \cdots, x_n$ have assigned values. We did not give an exposition of the method for the characterization of correlation§ by the equation of a

* Persons, W. M., "Correlation of Time Series", Handbook of Mathematical Statistics, (1924), pp. 150–65.

† Yule, G. U., Jour. Roy. Stat. Soc., vol. 60 (1897), p. 812.

Rietz, H. L., "Mathematical Statistics", The Carus Mathematical Monograph, No. 3, (1927), pp. 92–98.

Rietz, H. L., and Others: Handbook of Mathematical Statistics, (1924), pp. 139–146

‡ Rietz, H. L., and Others: Handbook of Mathematical Statistics, (1924), pp. 139–146.

§ Pearson, Karl, "On a general method of determining the successive terms in a skew regression line", Biometrica, vol. 13 (1920–21), pp. 296–300.

curve of regression that passes approximately through the means of arrays of a given system of variables, and we have not discussed the theory* of non-linear regression in n variables. The correlation surface method† of characterizing correlation has not been touched, and we did not discuss the properties of normally correlated‡ distributions. We made no statement concerning Bachelier's fundamental§ work on the theory of correlation in which he treats continuous probabilities of two or more variables. The Gram-Charlier system of representation of distributions|| has not been discussed. We did not give an exposition of the progress that has been made by Narumi¶, Pearson*, and Camp† in solving the problem of passing from the Pearsonian system of generalized frequency curves to analogous surfaces for the characterization of the distribution of two correlated variables.

---

* Biometrica, vol. 10 (1914–15), pp. 393–411.

Proc. Roy. Soc., A, vol. 91 (1915), pp. 494–498.

Rietz, H. L., "Report on statistics", Bull. Amer. Math. Soc., vol. 30 (1924), pp. 437–439.

Rietz, H. L., "Mathematical Statistics", The Carus Mathematical Monograph, No. 3 (1927), pp. 92–98.

† Rietz, H. L., "Mathematical Statistics", The Carus Mathematical Monograph, No. 3 (1927), pp. 104–108.

Whittaker & Robinson, "The Calculus of Observations" (1924), pp. 317–342.

‡ Rietz, H. L., "Mathematical Statistics", The Carus Mathematical Monograph, No. 3 (1927), pp. 108–112.

Rietz, H. L., "On the theory of correlation with special reference to certain significant loci on the plane of distribution", Ann. of Math. (2), vol. 13 (1912), pp. 195–196.

Pearson, Karl, "On the theory of contingency and its relation to association and normal correlation", Drapers' Company Research Memoirs (Biometric Series I) (1904), p. 10.

Czuber, E., "Theorie der Beobachtungsfehler" (1891), pp. 355–382.

McMahon, J., "Hyperspherical goniometry and its application to correlation theory for n variables", Biometrica, vol. 15 (1923), pp. 192–208.

§ Bachelier, Louis, "Calcul des Probabilités" (1912), Chapters 17 & 18.

|| Pearson, Karl, "On a general method of determining the successive terms in a skew regression line", Biometrica, vol. 13 (1920–21), pp. 296–300.

Rietz, H. L., and Others: Handbook of Mathematical Statistics (1924), pp. 114–116.

Bowley (London, 1920), Chapter 3, Part. 2.

Edgeworth, "The law of error", Camb. Phil. Trans., vol. 20 (1904).

Charlier, "Über das Fehlergesetz" and "Die zweite Form des Fehlergesetzes", Arkiv för Mat., Astr. och Fys., Bd. 2, No. 8.

¶ Narumi, Seimatsu, "On the general form of bivariate frequency distributions which are mathematically possible when regression and variation are subjected to limiting conditions", Biometrica, vol. 15 (1923), pp. 77–88, pp. 209–221.

* Pearson, Karl, "Notes on skew frequency surfaces", Biometrica, vol. 15 (1923), pp. 222–244.

† Camp, B. H., "Mutually consistent multiple regression surfaces", Biometrica, vol. 17 (1925), pp. 443–458.

We have made no mention of the correlation between functions, continuous variables,* and spurious† correlation.

The above mentioned partial and total omissions of concepts belonging to the English school have been made because (1) most of the material is so available elsewhere, and (2) it is our intention to limit the paper to the problem of two variables.

The following concepts belonging to the Italian school have been omitted because an exposition of them is somewhat beyond the scope of this paper.

Pietra,‡ in a recent paper, refers to an unpublished paper by Gini in which Gini develops an expression for the *Index of Systematic Connection*; that is, the index of the connection which may be expected when the number of observations being infinite, the influence of chance could be considered as eliminated.

In another unpublished paper by Gini, mentioned by Pietra,§ Gini establishes *indices of correlation between series of groups.* The groups considered are as follows: (a) Each group is composed of a given number, the same number for all groups, of heterogeneous quantities, to each one of which corresponds a homogeneous quantity of the other group; (b) The groups are composed of a number of homogeneous quantities, each one of which may be compared indifferently with any of the corresponding

---

* Yasukawa, Kazutaro, "On the means, standard deviations, correlations, and frequency distributions of functions of variables", Biometrica, vol. 17 (1925), pp. 211–238.

Rider, P. R., "The correlation between two variates one of which is normally distributed etc.", Amer. Math. Monthly, vol. 31 (1924), pp. 227–232.

Rietz, H. L., "Frequency distributions obtained by certain transformations of normally distributed variates", Ann. of Math. (2), vol. 23 (1921–22), pp. 292–300.

Huntington, E. V., "Mathematics and statistics etc.", Amer. Math. Monthly, vol. 26 (1919), pp. 421–434.

Rietz, H. L., Quart. Publ. Amer. Statistical Assoc., vol. 16 (1919), pp. 472–476.

Pearson, Karl, On the correlation between two variates", Amer. Math. Monthly, vol. 32, pp. 70–73.

Weida, F. M., "On the correlation between two functions", Amer. Math. Monthly, vol. 33, pp. 440–444.

† Pearson, Karl, "On a form of spurious correlation that may arise when indices are used in the measurement of organs", Proc. Roy. Soc., vol. 60 (1897), p. 489.

Yule, G. U., "On the interpretations of correlations between indices or ratios", Jour. Roy. Stat. Soc., vol. 73 (1910), p. 644.

Pearson, Karl, Phil. Trans., vol. 192 (1899), p. 257.

Yule, G. U., "Introduction to the Theory of Statistics" (1922), pp. 210–225.

Yule, G. U., "Why do we sometimes get nonsense correlations between time series?—A study in sampling and the nature of time series", Jour. Roy. Stat. Soc., vol. 89 (1926), Pt. I, pp. 1–64.

‡ Metron, vol. 4, Nos. 3–4 (1925), pp. 433–440.

§ Metron, vol. 4, Nos. 3–4 (1925), pp. 444–450.

group of the other series. An example of (a) is the study of the concordance between the annual prices of n commodities in two different cities. An example of (b) is the study of the concordance between parents and sons with reference to stature.

In a recent paper, Pietra* has extended the theory of dissimilarity, connection, and concordance, as developed by Gini, to cyclical series—series which present modalities according to a natural order of succession in which we can not find two modalities that may be considered as the first one and the last one of the succession, except in the case of a particular agreement. An example is the series of weddings according to the days of the week.

PART III.

The theory of correlation as an integral part of the theory of probability†.

16. In order to make the theory of correlation an integral part of the theory of probability, Tschuprow introduced the concept of *stochastic dependence*; namely,

Two variables x and y are said to be *stochastically connected* if for every x_j $(j = 1, 2, \cdots, m)$ of a *non-contingent* x, there exists a *contingent* y. A variable y which assumes, with definite probabilities, i different values is said to be a *contingent variable of the i-th order* and the totality of the i values and their respective probabilities is said to be the *law of distribution of the contingent variable.*

To illustrate: Suppose that we are throwing two dice. Let x be the number thrown with the first die and let y be the sum thrown with the two dice. Then if x is known to be, say 3, y would take the values 4, 5, 6, 7, 8, and 9 with definite probabilities. In this case, y is said to be *stochastically connected* with x. Also, x is a *non-contingent* variable and y is a *contingent* variable of the 6-th order.

If x is contingent, y is also contingent provided first, y is observed without reference to the value of x; and secondly, y and x are functionally connected. In general, a variable y and a set of variables x_k $(k = 1, 2, \cdots, n)$ are stochastically connected if, for a fixed x_1, for a fixed x_2, $\cdots$, and for a fixed x_n, there exists a contingent y. Again, a variable y and a set of variables x_k $(k = 1, 2, \cdots, n)$, are stochastically connected if for at least $(n-1)$ fixed x_k's, there exists a contingent y. Also, if at least one of the x_k's is contingent and if y and the set x_k $(k = 1, 2, \cdots, n)$, are

* Metron, vol. 4. Nos. 3-4 (1925), pp. 450-558.

† A. A. Tschuprow, "Grundbegriffe und Grundprobleme der Korrelationstheorie", Berlin (1925).

functionally connected; then y too is contingent provided y is observed without reference to the value of one of the contingent x_k's.

The fundamental notion of statistical correlation theory for stochastically connected variables is dependent upon the notion of contingent variables while the fundamental notion of statistical correlation theory for functionally connected variables is dependent upon the common notion of real variables.

The clear differentiation of the meaning of stochastic connection and functional connection is the first step necessary for the understanding of the character of the research in statistical correlation in its differentiation from the character of the research in the laws pertaining to the natural sciences. Stochastically connected variables may reach the investigator in the same ways as do the individual contingent variables. In a functional connection, the non-contingent variables may be changed to stochastically connected contingent variables through the errors of observations connected with the measurements.

When we have to work with a contingent variable, we have everything which may be said about it, in case the law of distribution is fixed. The determination of all the possible values of the variables and their probabilities, therefore, appears to be the main task of the investigation. In this form, however, our knowledge about the contingent variable is complete but can not be handled. One distribution law is hard to compare with another distribution law unless both of them are very simple. We must, therefore, rely upon suitably constructed index numbers for the comparison of distributions.

17. **Mathematical expectation and mean error.** Let $x_1, x_2, \cdots, x_k$ be the k values af a kth order contingent variable and let $p_1, p_2, \cdots, p_k$ be the respective probabilities of the values $x_1, x_2, \cdots, x_k$. Then the *mathematical expectation* e_x, of x, is defined by

$$e_x = \sum_1^k {}_j p_j x_j. \tag{57}$$

The *mean error* σ_x, of x, is defined by

$$\sigma_x = \left\{ \sum_1^k {}_j p_j (x_j - e_x)^2 \right\}^{1/2}. \tag{58}$$

The square of the mean error is said to be the *spread*. Hence, we say that σ_x^2 is the *spread* of x.

18. To introduce certain fundamental notions we let x represent a variable which is defined for the set of values $x_j\,(j = 1, 2, \cdots, m)$ and we let y represent a variable which is defined for the set of values $y_i\,(i = 1, 2, \cdots, n)$. Let $p_j\,(j = 1, 2, \cdots, m)$ represent the probabilities assigned respectively

to the x_j and let $p_i (i = 1, 2, \cdots, n)$ represent the probabilities assigned respectively to the y_i. The totality of the possible y_i's of y, for a particular x_j of x, together with the p_i probabilities assigned respectively to the y_i's we shall call the *limited law of distribution* of y for the particular x_j of x. The values of the mathematical expectation e_y of y, the mean error σ_y of y, and the spread σ_y^2 of y, computed on the principle of the limited law of distribution, we shall call, respectively, the *limited mathematical expectation*, the *limited mean error*, and the *limited spread*. The analytic expression defining the limited mathematical expectation of y as a function of the value x_j of x, is said to be the *Regression Equation* of y on x, or *Regression Line*. Also, the analytic expressions defining the limited mean error of y, and the limited mean spread of y as functions of the corresponding value of x are known as *Skedastic Equations* or *Skedastic Lines*.

If the limited mean error of y remains constant for all possible values of x, the connection of y with x is said to be *homoskedastic*; and if the limited mean error of y does not remain constant for all possible values of x, the connection of y with x is said to be *heteroskedastic*.

19. **Stochastic independence.** Stochastic independence forms one of the corner stones in statistical correlation. A variable y is said to be stochastically independent of another variable x, if the limited law of distribution of y for all possible values of x remains the same. It is comparatively easy to prove that if y is stochastically independent of x, then x is stochastically independent of y. Hence, two contingent variables are independent of each other.

20. **The law of correlation and non-correlation.** It appears that Tschuprow does not give Pearson as full credit as Pearson deserves in connection with the conceptions of correlation. Tschuprow would say, that a variable y is correlated with a variable x, if the limited mathematical expectation of y varies with x; and a variable y is not correlated with a variable x, if the limited mathematical expectation of y remains constant for all possible values of x. In other words, y is correlated with x whenever the theoretical distributions in arrays of y's are not identical for all possible assigned values of x; and y is uncorrelated with x whenever the theoretical distributions in arrays of y's are identical with each other for all possible values of x. By the identity of the theoretical distributions in arrays of y's, it is meant that they should have equal means, standard deviations, and parameters required to completely fix the distributions. This concept appears to be a somewhat more positive and precise conception than has been current among English statisticians. Still, it seems fairly obvious that Pearson held nearly this view when he wrote his con-

ception quoted in part I of this paper from his Draper's Company Research Memoir.

If the regression line of y on x is parallel to the x axis, then y is not correlated with x; but if the regression line is not parallel to the x axis, then y is correlated with x.

The *law of non-correlation* is of great value in the study of statistical correlation. It differs from the law of stochastic independence in that if a variable y is stochastically independent of x, y can not be correlated with x. Conversely, if y is not correlated with x, it does not follow that y is stochastically independent of x; namely, the limited mathematical expectation of y might remain constant for all values of x, but the limited law of distribution might be different for different values of x. Again, if y is not correlated with x, it does not follow that x is not correlated with y.

21. **The a-priori law of dependence.** Let x be an m-th order contingent variable and let y be an n-th order contingent variable. Let $p_{j|}$ be the probability that $x = x_j$ $(j = 1, 2, \cdots, m)$, let $p_{|i}$ be the probability that $y = y_i$ $(i = 1, 2, \cdots, n)$, let $p_{j|i}$ be the probability that $x = x_j$ and $y = y_i$ at the same time, let $p_{|i}^{(j)}$ be the probability that, when $x = x_j$, $y = y_i$, let $p_{j|}^{(i)}$ be the probability that, when $y = y_i$, $x = x_j$. Then we may write that

$$p_{j|i} = p_{j|}\, p_{|i}^{(j)} \quad \text{and} \quad p_{j|i} = p_{|i}\, p_{j|}^{(i)}. \tag{59}$$

According to article 19 of this paper, y is stochastically independent of x if $p_{|i}^{(j)} = p_{|i}$ $(i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, m)$. Conversely, if $p_{|i}^{(j)} = p_{|i}$ for every possible value of i and for every possible value of j, then the limited law of distribution of y remains the same for all values of x and consequently y is independent of x. If in $p_{j|}\, p_{|i}^{(j)} = p_{|i}\, p_{j|}^{(i)}$ we put $p_{|i}^{(j)} = p_{|i}$, we find that $p_{j|}\, p_{|i} = p_{|i}\, p_{j|}^{(i)}$ and $p_{j|}^{(i)} = p_{j|}$. If $p_{|i}^{(j)} = p_{|i}$ for every i and j, then $p_{j|}^{(i)} = p_{j|}$ for every i and j. Thus y stochastically independent of x implies x stochastically independent of y.

Again, if x and y are mutually independent, then $p_{j|i} = p_{j|}\, p_{|i}$ for all possible values of i and j. Conversely, if $p_{j|i} = p_{j|}\, p_{|i}$ for all possible values of i and j, then $p_{|i}^{(j)} = p_{|i}$ for all possible values of i and j and x and y are mutually independent.

If y and x are mutually independent, then all differences $p_{j|i} - p_{j|}\, p_{|i} = 0$. If one or more of the differences $p_{j|i} - p_{j|}\, p_{|i}$ do not equal zero, then y and x are not independent.

22. **Pearson's mean square contingency "φ^2".** A measure of the departure from mutual independence is Pearson's "*Mean Square Contingency* φ^2" which is defined as

$$\varphi^2 = \sum_1^n{}_i \sum_1^m{}_j \frac{\{p_{j|i} - p_{j|} p_{|i}\}^2}{p_{j|} p_{|i}}. \tag{60}$$

If y and x are mutually independent, then all the differences $p_{j|i} - p_{j|} p_{|i}$ are zero and therefore $\varphi^2 = 0$. Conversely, $\varphi^2 = 0$, only when all the differences $p_{j|i} - p_{j|} p_{|i}$ are zero; and if $\varphi^2 = 0$, it follows at once that y and x are mutually independent.

23. The correlation coefficient. Another important group of procedures which is used in the examination of stochastically connected variables proceeds from the study of the limited law of distribution. If two variables y and x are stochastically independent, then the limited law of distribution of y remains the same for all values of x. If two variables y and x are *not* stochastically independent, then the limited law of distribution of y does *not* remain the same for all values of x. The law which defines how the various characteristics of the limited law of distribution of y change form as the variable x runs through all its possible values is, therefore, one of the characteristics of the law of dependence between y and x.

To study these changes orderly, we introduce a number of new symbols. Let $E_{f|g}$ represent the mathematical expectation of the f-th power of x and the g-th power of y, so that

$$E_{f|g} = \sum_1^n{}_i \sum_1^m{}_j p_{j|i}\, x_j^f\, y_i^g. \tag{61}$$

Let us also write that

$$U_{f|g} = \sum_1^n{}_i \sum_1^m{}_j p_{j|i} \{x_j - E_{1|0}\}^f \{y_i - E_{0|1}\}^g, \tag{62}$$

and

$$r_{f|g} = \frac{U_{f|g}}{U_{2|0}^{1/2f}\, U_{0|2}^{1/2g}}. \tag{63}$$

If $g = 0$, we obtain the corresponding parameter with which the law of distribution of x is characterized. If $f = 0$, we obtain the corresponding parameter with which the law of distribution of y is characterized. Thus $E_{1|0}$ is the mathematical expectation of x, $E_{0|1}$ is the mathematical expectation of y, $U_{2|0}$ is the spread of x, and $U_{0|2}$ is the spread of y.

The parameters $E_{f|g}$, $U_{f|g}$, and $r_{f|g}$ will be used continually. If we know the law of dependence, all the parameters can be determined; and the value of each parameter is unique. Conversely, the law of dependence can be uniquely determined, if the parameters $E_{f|g}$, $U_{f|g}$, and $r_{f|g}$ are known. A proof of these statements is somewhat beyond the scope of this paper. If the laws of dependence are discontinuous, the proof is not difficult; if

the laws of dependence are continuous, the proof is rather difficult. In the latter case we meet the "Problem of Moments".

Of the three parameter systems, the system $r_{f|g}$ is conspicuous because each member of the system is an abstract number. This fact, makes the $r_{f|g}$ system particularly useful for the comparison of the laws of dependence. The first in the series of the $r_{f|g}$ system is the well known Pearsonian coefficient of correlation which is given by

$$r_{1|1} = \frac{U_{1|1}}{U_{2|0}^{1/2}\, U_{0|2}^{1/2}} = \frac{U_{1|1}}{\sigma_x\, \sigma_y}. \tag{64}$$

It is easily shown that the absolute value of the correlation coefficient can not be greater than unity. Since the mathematical expectation of a variable, which is not defined for negative values, can never be negative, we have that

$$\sum \left\{ \frac{x - E_{1|0}}{\sigma_x} - r_{1|1} \frac{y - E_{0|1}}{\sigma_y} \right\}^2 \geqq 0.$$

But

$$\sum \frac{\{x - E_{1|0}\}^2}{\sigma_x^2} = \sum \frac{\{y - E_{0|1}\}^2}{\sigma_y^2} = 1 \quad \text{and} \quad \sum \frac{\{x - E_{1|0}\}\{y - E_{0|1}\}}{\sigma_x\, \sigma_y} = r_{1|1}.$$

Hence,

$$\sum \left\{ \frac{x - E_{1|0}}{\sigma_x} - r_{1|1} \frac{y - E_{0|1}}{\sigma_y} \right\}^2 = 1 - r_{1|1}^2;$$

thus

$$1 - r_{1|1}^2 \geqq 0 \quad \text{and} \quad r_{1|1}^2 \leqq 1.$$

If $E_{|1}^{(j)}$ is the limited mathematical expectation of y for $x = x_j$, and if $U_{|2}^{(j)}$ is the limited spread of y for $x = x_j$, then

$$E_{|1}^{(j)} = \sum_1^n{}_i\, p_{|i}^{(j)}\, y_i, \qquad U_{|2}^{(j)} = \sum_1^n{}_i\, p_{|i}^{(j)}\, \{y_i - E_{|1}^{(j)}\}^2.$$

Likewise, we have that

$$E_{1|}^{(i)} = \sum_1^m{}_j\, p_{j|}^{(i)}\, x_j, \qquad U_{2|}^{(i)} = \sum_1^m{}_j\, p_{j|}^{(i)}\, \{x_j - E_{1|}^{(i)}\}^2.$$

From our definitions, we have that

$$E_{0|1} = \sum_1^m{}_j\, p_{j|}\, E_{|1}^{(j)}, \qquad E_{1|0} = \sum_1^n{}_i\, p_{|i}\, E_{1|}^{(i)},$$

$$\sum_1^m{}_j\, p_{j|}\, U_{|2}^{(j)} = \sum_1^n{}_i \sum_1^m{}_j\, p_{j|}\, p_{|i}^{(j)}\, \{(y_i - E_{0|1}) - (E_{|1}^{(j)} - E_{0|1})\}^2$$

$$= U_{0|2} - \sum_1^m{}_j\, p_{j|}\, (E_{|1}^{(j)} - E_{0|1})^2,$$

$$\sum_1^n {}_i p_{|i}\, U_{2|0}^{(i)} = U_{2|0} - \sum_1^n {}_i p_{i|}\, (E_{1|}^{(i)} - E_{1|0})^2.$$

If the variables are mutually independent, then for any h, $E_{h|}^{(i)} = E_{h|0}$, $E_{|h}^{(j)} = E_{0|h}$, for $j = 1, 2, \cdots, m$ and $i = 1, 2, \cdots, n$.

It is easily shown that three necessary and sufficient conditions for the mutual independence of two contingent variables x and y are that

$$E_{f|g} = E_{f|0}\, E_{0|g}, \quad U_{f|g} = U_{f|0}\, U_{0|g}, \quad r_{f|g} = r_{f|0}\, r_{0|g}, \tag{65}$$

for all positive values of f and g.

24. **Regression equation.** We shall now study the limited mathematical expectation. The equation which defines the connection between the mathematical expectation of y and the corresponding values of x is called the *regression equation* of y on x:

$E_{|1}^{(j)} = f(x_j)$ is the *regression equation* of y on x and

$E_{1|}^{(i)} = f(y_i)$ is the *regression equation* of x on y.

If the regression of y on x assumes the form of a parabola given by the equation

$$E_{|1}^{(j)} = a_{|0} + a_{|1}\, x_j + a_{|2}\, x_j^2 + \cdots + a_{|f} x_j^f, \tag{66}$$

it is readily seen that the coefficients $a_{|f}$ may be expressed in terms of the parameters E, U, and r. We need only multiply both sides of (66) by $p_{j|}\, x_j^h$ and form the j sum. Since $\sum_j p_j\, x_j^h\, E_{|1}^{(j)} = E_{h|1}$, we have that

$$E_{h|1} = a_{|0}\, E_{h|0} + a_{|1}\, E_{h+1|0} + \cdots + a_{|f}\, E_{h+f|0}. \tag{67}$$

If, in (67), we in turn give to h the values $0, 1, 2, \cdots, f$; we get $f+1$ linear equations which will suffice to determine the $f+1$ coefficients $a_{|f}$. Replacing the a's in (66) by the values just found, we obtain the f-th degree parabola regression equation desired.

In case we have linear regression expressed by the equation $E_{|1}^{(j)} = a_{|0} + a_{|1}\, x_j$ we find that

$$a_{|1} = \frac{E_{1|1} - E_{1|0}\, E_{0|1}}{E_{2|0} - E_{1|0}^2} \quad \text{and} \quad a_{|0} = E_{0|1} - a_{|1}\, E_{1|0}.$$

Hence, the regression equation becomes

$$E_{|1}^{(j)} - E_{0|1} = \frac{E_{1|1} - E_{1|0}\, E_{0|1}}{E_{2|0} - E_{1|0}^2}\, (x_j - E_{1|0}). \tag{68}$$

The conditions which must be fulfilled by the parameters E so that the regression of y on x is linear is that

$$(69) \qquad \frac{E_{h|1}-E_{h|0}\,E_{0|1}}{E_{h+1|0}-E_{h|0}\,E_{1|0}} = \frac{E_{1|1}-E_{1|0}\,E_{0|1}}{E_{2|0}-E_{1|0}^2}$$

for all values of $h>0$.

The regression equation may also be written in the form

$$(70) \quad E_{|1}^{(j)}-E_{0|1} = b_{|0}+b_{|1}(x_j-E_{1|0})+b_{|2}(x_j-E_{1|0})^2+\cdots+b_{|f}(x_j-E_{1|0})^f.$$

If we multiply both sides of (70) by $p_{j|}\,(x_j-E_{1|0})^h$ and form the j sum, noticing that $\sum_j p_{j|}\,(x_j-E_{1|0})^h\,(E_{|1}^{(j)}-E_{0|1}) = U_{h|1}$, we get

$$(71) \qquad U_{h|1} = b_{|0}\,U_{h|0}+b_{|1}\,U_{h+1|0}+b_{|2}\,U_{h+2|0}+\cdots+b_{|f}\,U_{h+f|0}.$$

If we let h take the values $0, 1, 2, \cdots, f$; we obtain a system of $f+1$ equations from which we may determine the values of the b's. The substitution of these values of the b's in (70) will give us the desired regression equation.

The conditions for linear regression are

$$(72) \quad E_{|1}^{(j)}-E_{0|1} = \frac{U_{1|1}}{U_{2|0}}(x_j-E_{1|0}) \quad \text{and} \quad E_{1|}^{(i)}-E_{1|0} = \frac{U_{1|1}}{U_{0|2}}(y_i-E_{0|1}).$$

The coefficients $U_{1|1}/U_{2|0} = (\sigma_y/\sigma_x)\,r_{1|1}$ and $U_{1|1}/U_{0|2} = (\sigma_x/\sigma_y)\,r_{1|1}$ are known as the *regression coefficients.*

If we adopt the so-called "*Normal Coördinates*" (coördinates obtained by dividing the left and right hand deviations from the expectation by the spread) we may write for the equation of linear regression that

$$(73) \quad Y_{|1}^{(j)} = r_{1|1}\,X_j, \text{ where } \frac{E_{|1}^{(j)}-E_{0|1}}{\sigma_y} = Y_{|1}^{(j)} \quad \text{and} \quad \frac{x_j-E_{1|0}}{\sigma_x} = X_j.$$

The condition for linear regression is

$$(74) \qquad r_{h|1} = r_{1|1}\,r_{h+1|0} \quad \text{for } h = 2, 3, 4, \cdots.$$

Similarly, we may write for the second degree regression equation that

$$(75) \quad Y_{|1}^{(j)} = \frac{r_{3|0}\,r_{1|1}-r_{2|1}}{r_{4|0}-r_{3|0}^2-1}+\left\{r_{1|1}-\frac{r_{3|0}\,(r_{2|1}-r_{3|0}\,r_{1|1})}{r_{4|0}-r_{3|0}^2-1}\right\}X_j+\frac{r_{2|1}-r_{3|0}\,r_{1|1}}{r_{4|0}-r_{3|0}^2-1}X_j^2.$$

We observe, that in this case, the member free from X_j in (75) does not disappear and if $r_{1|1}=0$, then (75) becomes

$$(76) \qquad Y_{|1}^{(j)} = \frac{r_{2|1}}{r_{4|0}-r_{3|0}^2-1}\left\{X_j^2-r_{4|0}\,X_j-1\right\}.$$

The condition for second degree regression is

$$(77)\qquad \frac{r_{h|1} - r_{1|1}\, r_{h+1|0}}{r_{h+2|0} - r_{h+1|0}\, r_{3|0} - r_{h|0}} = \frac{r_{2|1} - r_{1|1}\, r_{3|0}}{r_{4|0} - r_{3|0}^2 - 1} \quad \text{for} \quad h = 3, 4, 5, \cdots.$$

If y and x are not correlated, the regression equation of y on x is $Y_{|1}^{(j)} = 0$. Also, in order that y and x are not correlated, the parameters E, U, and r must be such that $E_{h|1} = E_{h|0}\, E_{0|1}$, $U_{h|1} = 0$, $r_{h|1} = 0$, for $h = 1, 2, 3, \cdots$.

If in $r_{h|1} = 0$, we put $h = 1$, then $r_{1|1} = 0$. If y is not correlated with x, then $r_{1|1} = 0$. Conversely, if $r_{1|1} = 0$, it does not follow that y is not correlated with x. If $r_{1|1} = 0$, and the regression is linear, then y is not correlated with x, and $E_{|1}^{(j)} = E_{0|1}$. If the regression is not linear and $r_{1|1} = 0$, it does not follow that $E_{|1}^{(j)} = E_{0|1}$.

It is important to notice, that if $U_{h|1} = 0$ for $h = 1, 2, \cdots$, it does not follow that $U_{1|h} = 0$ for $h = 2, 3, \cdots$: If y is not correlated with x, it does not follow that x is not correlated with y. If there is no correlation, it does not follow that there is independence—not even in the case where there is no correlation of y on x as well as of x on y: If $U_{h|1} = 0$ and $U_{1|h} = 0$ for $h = 1, 2, 3, \cdots$ it does not follow that $U_{f|g} = U_{f|0}\, U_{0|g}$ when f as well as g are different from unity.

25. **The correlation ratio.** Another group of procedures, whose aim is to represent the laws of dependence, is the computation of numbers which are similar to the mean square contingency. These numbers differ from the mean square contingency in that they not only use the probabilities of the possible values of the variables, but also use the possible values. Two such numbers must be studied in more detail: The correlation coefficient and the correlation ratio. The computation of the correlation coefficient may be considered at present as the most favored of all procedures that may be used for the examination of two stochastically connected variables. Since the correlation coefficient equals the correlation ratio under certain assumptions, we will study the correlation ratio first.

The correlation ratio, discovered by Pearson, is defined as*

$$(78)\qquad \eta_{y|x}^2 = 1 - \frac{1}{U_{0|2}} \sum_1^m {}_j p_{j|}\, U_{|2}^{(j)}.$$

From (78) we see that the correlation ratio is the complement of the quotient of the average limited spread divided by the non-limited total spread. Since, we have seen that

* Karl Pearson, "On the general theory of skew correlation etc.", Drapers' Company research memoirs, Biometric Series, II (1905).

$$\sum_j p_{j|}\, U_{|2}^{(j)} = U_{0|2} - \sum_j p_{j|}\, \{E_{|1}^{(j)} - E_{0|1}\}^2,$$

it follows from our definition that

$$\eta_{y|x}^2 = \frac{1}{U_{0|2}} \sum_j p_{j|}\, (E_{|1}^{(j)} - E_{0|1})^2. \tag{79}$$

The stochastical connection of the variable y with the variable x consists in that y, even when the value of x is determined, remains a contingent variable. In other words, the limited spreads of y remain different from zero. If y is in functional connection with x, then all the limited spreads of y are zero and the correlation ratio of y on x is equal to unity. Conversely, if the correlation ratio of y on x is equal to unity, then the average limited spread of y is zero. This is possible when all the individual limited spreads of y are zero and thus y is in functional connection with x. Since the average limited spread is always positive, the correlation ratio can never be greater than unity. Thus, the value unity for the correlation ratio indicates that y and x are functionally connected. Again, the minimum value of the correlation ratio is zero, since the average of the limited spreads of y can not be greater than the total non-limited spread. In a similar fashion, one may discuss the correlation ratio of x on y.

To compute the correlation ratio, we may use (78) as well as

$$\eta_{y|x}^2 = \sum_j p_{j|}\, \{Y_{|1}^{(j)}\}^2. \tag{80}$$

If the regression equation is known, the value of $Y_{|1}^{(j)}$ is readily obtained. In case we have the second degree regression equation (75), if we make certain transformations, we find that

$$\eta_{y|x}^2 = r_{1|1}^2 + \frac{(r_{2|1} - r_{3|0}\, r_{1|1})^2}{r_{4|0} - r_{3|0}^2 - 1}. \tag{81}$$

If the regression of y on x is linear, with equation $E_{|1}^{(j)} = r_{1|1}\, X_j$ we obtain, since $\sum_j p_{j|}\, X_j^2 = 1$, that $\eta_{y|x}^2 = r_{1|1}^2$. Hence when the regression is linear, the correlation coefficient and the correlation ratio are the same absolute number. If the regression is not linear, then $r_{1|1}^2 < \eta_{y|x}^2$. This is easily shown. Consider $y_{|1}^{(j)} = E_{0|1} + (U_{1|1}/U_{2|0})\,(x_j - E_{1|0})$. Since $\sum_j p_{j|}\, (E_{|1}^{(j)} - y_{|1}^{(j)})^2 = U_{0|2}\, (\eta_{y|x}^2 - r_{1|1}^2)$, we find that

$$\eta_{y|x}^2 - r_{1|1}^2 = \frac{1}{U_{0|2}} \sum_j p_{j|}\, (E_{|1}^{(j)} - y_{|1}^{(j)})^2. \tag{82}$$

The difference $\eta_{y|x}^2 - r_{1|1}^2$ can not be negative. It can be zero, only when all the $E_{|1}^{(j)}$ coincide with the corresponding $y_{|1}^{(j)}$. In other words, the line

$$(83) \qquad y_{|1}^{(x)} = E_{0|1} + \frac{U_{1|1}}{U_{2|0}}(x_j - E_{1|0})$$

is the true regression line.

26. **Normal correlation.** A particularly favored position belongs to the correlation coefficient when the stochastical connection has the form of a so-called "Normal Correlation". Under normal correlation we understand the case when the variables x and y are defined for all real positive and negative numbers. The probability that the value of x is between x and $x+dx$ and y is between y and $y+dy$ is given by

$$(84) \qquad \frac{1}{2\pi\{U_{2|0}\,U_{0|2} - U_{1|1}^2\}^{1/2}}\, e^{-\frac{U_{0|2}[x-E_{1|0}]^2 - 2U_{1|1}[x-E_{1|0}][y-E_{0|1}] + U_{2|0}[y-E_{0|1}]^2}{2[U_{2|0}\,U_{0|2} - U_{1|1}]}}\, dx\,dy.$$

If we put

$$\mathfrak{x} = \frac{x - E_{1|0}}{\{U_{2|0}\}^{1/2}} \qquad \text{and} \qquad \mathfrak{y} = \frac{y - E_{0|1}}{\{U_{0|2}\}^{1/2}},$$

then the probability that $\mathfrak{x}$ is between $\mathfrak{x}$ and $\mathfrak{x}+d\mathfrak{x}$ and $\mathfrak{y}$ is between $\mathfrak{y}$ and $\mathfrak{y}+d\mathfrak{y}$ is given by

$$(85) \qquad \frac{1}{2\pi\{1-r_{1|1}^2\}^{1/2}} e^{-\frac{\mathfrak{x}^2 - 2r_{1|1}\mathfrak{x}\mathfrak{y} + \mathfrak{y}^2}{2\{1-r_{1|1}^2\}}}\, d\mathfrak{x}\, d\mathfrak{y}.$$

The dependence law is therefore determined in the case of normal correlation through the value of the correlation coefficient together with the values of both mathematical expectations and both spreads $E_{1|0}, U_{2|0}, E_{0|1}, U_{0|2}$ which characterize the law of distribution of x and the law of distribution of y and belong to the determination of the system of normal coördinates.

In the case of normal coördinates, the correlation coefficient appears to be the key to what we wish to know concerning the connection of the variables: if we know the value of $r_{1|1}$ the law of dependence is determined and all else may be deduced by formal mathematical operations.

Since the value of the correlation coefficient determines the law of dependence; in the case of normal correlation, all other measures may be represented as functions of $r_{1|1}$.

The law of dependence is symmetrical. The laws of distribution of x as well as y have the form of the Gauss-Laplace law of error. The regression of y on x as well as the regression of x on y is linear. The correlation coefficient and the two correlation ratios are equal; consequently, the correlation ceefficient may be considered a dependable measure of the degree of connection.

All the limited laws of distribution of x as well as y have the form of the Gauss-Laplace law of error. The dependence of y on x and of x on y

is homoskedastic. The limited spread of x remains constant for all values of y and has the value $U_{2|}^{(y)} = U_{2|0}\{1-r_{1|1}^2\}$. The limited spread of y retains the same value for all values of x and is equal to $U_{|2}^{(x)} = U_{0|2}\{1-r_{1|1}^2\}$.

It is of interest to note that, in the case of normal correlation, the mean square contingency is connected with the correlation coefficient in that

$$\varphi^2 = \frac{r_{1|1}^2}{1-r_{1|1}^2}.$$

The normal correlation has a favored position because (1) the mathematical analysis is much simpler and more elegant in the case of normal correlation than otherwise; (2) out of the general formulas, by adopting normal correlation, all higher r-parameters are eliminated since they can be represented as functions of the correlation coefficient; (3) the formulas not only gain clearness, but are also practical.

To summarize, we have called attention to the fact that there are many conceptions of correlation. To measure the degree of correlation or association of statistical variables, the English shool of statisticians has introduced the correlation coefficient, correlation ratio, correlation index, tetrachoric correlation coefficient, biserial correlation coefficient, coefficient of mean square contingency, coefficient of association, coefficient of colligation, and the coefficient to measure correlation in rank. Gini has put forward a somewhat more general set of coefficients after separating the problems into those of connection and concordance. Tschuprow has suggested the possible place of a higher product moment formula; still his main idea was to conceive of correlation as belonging to a-priori probability. His work surely tends to make the theory more attractive to the mathematician.

LEHIGH UNIVERSITY, BETHLEHEM, PENNSYLVANIA.

KERNELS OF POSITIVE TYPE.*

By Caroline Seely.

In an earlier paper† I have proved the following theorem:

Theorem I. *If a kernel $K(s, t)$, with more than one characteristic constant, is such that for every positive number n there exists an iterated kernel $K^m(s, t)$, $m > n$, that is of positive type‡ with respect to the set of characteristic functions $\varphi_i(s)$, $\psi_i(t)$ of $K(s, t)$, then*

$$(1) \qquad \int_a^b \varphi_i(s)\,\varphi_j(s)\,ds = 0, \qquad \int_a^b \psi_i(s)\,\psi_j(s)\,ds = 0, \qquad (\lambda_i \neq \lambda_j),$$

that is, the $\varphi_i(s)$ and $\psi_i(s)$ each form an orthogonal set.

A function satisfying the hypothesis of this theorem will be said to satisfy Condition A.

In the present paper we shall study the effect of a condition analogous to Condition A in which the principal functions§ take the place of the characteristic, or, as we shall call them, following Goursat, fundamental functions of the kernel.

Let $k(s, t)$ be a canonical kernel of $K(s, t)$, corresponding to a characteristic constant c and φ_i, ψ_i the corresponding principal functions. Then‖

$$(2) \qquad k(s, t) = \varphi_1(s)\,\psi_1(t) + \varphi_2(s)\,\psi_2(t) + \cdots + \varphi_p(s)\,\psi_p(t),$$

where

$$(3) \qquad c\int_a^b k(s, t)\,\varphi_1(t)\,dt = \varphi_1(s), \qquad c\int_a^b k(s, t)\,\varphi_2(t)\,dt = \varphi_1(s) + \varphi_2(s),$$
$$\cdots, \quad c\int_a^b k(s, t)\,\varphi_p(t)\,dt = \varphi_{p-1}(s) + \varphi_p(s).$$

Here φ_1 is orthogonal to all the ψ_i except ψ_1, φ_2 to all the ψ_i except ψ_1 and ψ_2, $\cdots$, φ_p is orthogonal to all the ψ_i except ψ_{p-1} and ψ_p; also

$$(4) \qquad c\int_a^b \varphi_1(s)\,\psi_1(s)\,ds = 1, \quad c\int_a^b \varphi_2(s)\,\psi_1(s)\,ds = 1, \quad c\int_a^b \varphi_2(s)\,\psi_2(s)\,ds = 1,$$
$$\cdots, \quad c\int_a^b \varphi_p(s)\,\psi_{p-1}(s)\,ds = c\int_a^b \varphi_p(s)\,\psi_p(s)\,ds = 1.$$

---

* Received October 11, 1927.

† Annals of Math., vol. 20, pp. 172–176.

‡ Throughout this paper the term "positive" does not mean "positive definite"; the integral may be $\geqq 0$ but not < 0.

§ Goursat, Cours d'Analyse, vol. 3 (1923), pp. 405 ff. The characteristic (fundamental) functions are the solutions of the homogeneous Fredholm equations.

‖ Goursat, loc. cit., pp. 409, 410.

$\varphi_i(s)$ is a fundamental function of $K(s, t)$ and $\psi_p(s)$ a fundamental function of $K(t, s)$ for the characteristic constant c.

Goursat* gives the following expression for the $(n+1)$th iterated kernel of $k(s, t)$:

$$
\begin{aligned}
c^n\, k^{n+1}(s,t) = {} & \varphi_1(s)\{\psi_1(t) + C_n^1\,\psi_2(t) + C_n^2\,\psi_3(t) + \cdots\} + \cdots \\
& + \varphi_i(s)\{\psi_i(t) + C_n^1\,\psi_{i+1}(t) + C_n^2\,\psi_{i+2}(t) + \cdots\} + \cdots \\
& + \varphi_p(s)\,\psi_p(t),
\end{aligned} \tag{5}
$$

where $C_n^h = n!/(n-h)!\,h!$

We will assume that $k^{n+1}(s, t)$ is of positive type with respect to all linear combinations $\alpha\varphi_1 + \beta\varphi_2$ of the principal functions φ_1 and φ_2 of (3) and (4), where the α and β are real. It is assumed throughout this paper that all functions are of summable square, and that all multiple integrals occurring are independent of the order of integration.

Using the expression for $k^{n+1}(s, t)$ given by (5), we have

$$
\begin{aligned}
& \int_a^b\int_a^b k^{n+1}(s,t)\,\{\alpha\varphi_1(s) + \beta\varphi_2(s)\}\,\{\alpha\varphi_1(t) + \beta\varphi_2(t)\}\,ds\,dt \\
& = \frac{1}{c^n}\int_a^b\int_a^b \{\varphi_1(s)\,[\psi_1(t) + n\,\psi_2(t) + \cdots] \\
& \qquad + \varphi_2(s)\,[\psi_2(t) + n\,\psi_3(t) + \cdots] \\
& \qquad + \varphi_3(s)\,[\psi_3(t) + n\,\psi_4(t) + \cdots] + \cdots + \varphi_p(s)\,\psi_p(t)\} \\
& \qquad \times [\alpha\varphi_1(t) + \beta\varphi_2(t)]\,[\alpha\varphi_1(s) + \beta\varphi_2(s)]\,ds\,dt \qquad (6) \\
& = \frac{1}{c^{n+1}}\int_a^b [\alpha\varphi_1(s) + \beta((n+1)\,\varphi_1(s) + \varphi_2(s))]\,[\alpha\varphi_1(s) + \beta\varphi_2(s)]\,ds \\
& = \frac{1}{c^{n+1}}\Big[\alpha^2\int_a^b (\varphi_1(s))^2\,ds + \alpha\beta\Big\{(n+1)\int_a^b (\varphi_1(s))^2\,ds + 2\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds\Big\} \\
& \qquad + \beta^2\Big\{(n+1)\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds + \int_a^b (\varphi_2(s))^2\,ds\Big\}\Big] \geqq 0.
\end{aligned}
$$

This relation holds for all values of α and β; since it holds for $\beta = 0$, evidently $1/c^{n+1}$ is positive. Therefore the condition (6) reduces to the condition that the quadratic form in α and β be positive:

$$
\begin{aligned}
& \Big[(n+1)\int_a^b (\varphi_1(s))^2\,ds + 2\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds\Big]^2 \\
& \qquad \leqq 4\Big[\int_a^b (\varphi_1(s))^2\,ds\Big\{(n+1)\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds + \int_a^b (\varphi_2(s))^2\,ds\Big\}\Big],
\end{aligned} \tag{7}
$$

* Loc. cit., p. 411.

or

$$\left[(n+1)\int_a^b (\varphi_1(s))^2\,ds\right]^2 + 4(n+1)\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds\int_a^b (\varphi_1(s))^2\,ds + 4\left[\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds\right]^2$$
$$\leqq 4(n+1)\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds\int_a^b (\varphi_1(s))^2\,ds + 4\int_a^b (\varphi_1(s))^2\,ds\int_a^b (\varphi_2(s))d^2\,s.$$

Whence

$$\left[(n+1)\int (\varphi_1(s))^2\,ds\right]^2 \leqq 4\left\{\int_a^b (\varphi_1(s))^2\,ds\int_a^b (\varphi_2(s))^2\,ds - \left[\int_a^b \varphi_1(s)\,\varphi_2(s)\,ds\right]^2\right\} \leqq C \tag{8}$$

where C is a positive* constant, independent of n or is zero.

But $\int_a^b (\varphi_1(s))^2\,ds > 0$, since functions nearly everywhere zero are not admitted, as is usual in the theory. Therefore for sufficiently large n the left member of (8) can be made greater than any assigned constant, and we have a contradiction of (8).

On the other hand, if we assume that $k^{n+1}(s, t)$ is of positive type with respect to all linear combinations $\alpha\psi_p + \beta\psi_{p-1}$ of the principal functions ψ_p and ψ_{p-1}, we have, using the same expression for $k^{n+1}(s, t)$

$$\begin{aligned}\int_a^b\int_a^b k^{n+1}(t, s)\,\{\alpha\psi_p(s) + \beta\psi_{p-1}(s)\}\,\{\alpha\psi_p(t) + \beta\psi_{p-1}(t)\}\,ds\,dt \\ = (1/c^{n+1})\left[\alpha^2\int_a^b (\psi_p(s))^2\,ds\right. \\ + \alpha\beta\left\{(n+1)\int_a^b (\psi_p(s))^2\,ds + 2\int_a^b \psi_{p-1}(s)\,\psi_p(s)\,ds\right\} \\ \left. + \beta^2\left\{(n+1)\int_a^b \psi_{p-1}(s)\,\psi_p(s)\,ds + \int_a^b (\psi_{p-1}(s))^2\,ds\right\}\right] \geqq 0,\end{aligned} \tag{9}$$

whence, as in the earlier case,

$$\left[(n+1)\int_a^b (\psi_p(s))^2\,ds\right]^2 \leqq 4\left\{\int_a^b (\psi_p(s))^2\,ds\int_a^b (\psi_{p-1}(s))^2\,ds - \left[\int_a^b \psi_{p-1}(s)\,\psi_p(s)\,ds\right]^2\right\} \leqq C', \tag{10}$$

where C' is a constant independent of n, and (10) gives us the same contradiction as (8).

We have thus proved that a canonical kernel $k(s, t)$ of $K(s, t)$ cannot be such that an infinite number of its iterated kernels are of positive type with respect to all linear combinations of its principal functions unless $\varphi_2(s)$

* In virtue of the inequality of Schwartz.

and $\psi_{p-1}(s)$ do not exist and the set of its principal functions reduces to the pair of fundamental functions $\varphi_1(s)$ and $\psi_p(s)$. So we have the following theorem:

THEOREM II. *If every canonical kernel $k(s, t)$ of a kernel $K(s, t)$ is such that an infinite number of its iterated kernels are of positive type with respect to all linear combinations of its principal functions, then $K(s, t)$ has no principal functions except its fundamental functions, and all the poles of the resolvent kernel are simple.*

Since the principal functions $\varphi[\psi]$ of any canonical kernel $k(s, t)$ of $K(s, t)$ are orthogonal on the right [left] to $K(s, t) - k(s, t)$, and since the principal kernel of the nth iterated kernel $K^n(s, t)$ of $K(s, t)$ corresponding to the characteristic constant c^n of $K^n(s, t)$ is the nth iterated kernel of the principal kernel of $K(s, t)$ corresponding to c (or the sum of those corresponding to $+c$ and $-c$, if $K(s, t)$ has two equal characteristic constants of different signs)*, we see that Theorem II is equivalent to

THEOREM III. *If a kernel $K(s, t)$ is such that an infinite number of its iterated kernels $K^n(s, t)$ are of positive type with respect to all linear combinations of the principal functions φ and ψ of $K(s, t)$, then $K(s, t)$ has no principal functions other than its fundamental functions, and all the poles of the resolvent are simple.*

Functions satisfying the hypothesis of Theorem III will be said to satisfy Conditions B. Evidently Condition B includes Condition A, so that the fundamental functions Φ and Ψ of kernels satisfying Condition B form two sets of orthogonal functions as well as a single set of biorthogonal functions.

It may be asked whether Condition B is necessary in order that a kernel $K(s, t)$ be itself of positive type with respect to all functions of integrable square. We shall now show that this condition is not necessary, by giving an example of a function of positive type with respect to any function developable in a Fourier series convergent in the mean, and yet having principal functions that are not fundamental functions.

EXAMPLE. Consider, in the interval $(0, 2\pi)$, the function

$$(11) \quad K(s, t) = (2\sin s)\left(\frac{\sin t}{2} + \frac{\sin 2t}{4}\right) + (\sin s + 2\sin 2s)\left(\frac{\sin 2t}{2}\right).$$

Here $\varphi_1(s) = 2\sin s$ and $\psi_2(t) = (\sin 2t)/2$ are the fundamental functions corresponding to the characteristic constant $1/\pi$, and $\psi_1 = (\sin t)/2 + (\sin 2t)/4$ and $\varphi_2 = (\sin s + 2\sin 2s)$ are principal functions not fundamental functions. For

* See Goursat, loc. cit., p. 413. Note two obvious misprints: in line 14 from the bottom, for $K(x, y)$ read $K^n(x, y)$; in line 11 from the bottom, for $K^n(x, y)$ read $K(x, y)$.

$$(1/\pi)\int_0^{2\pi} K(s,t)(2\sin t)\,dt = 2\sin s = \varphi_1(s);$$

$$(1/\pi)\int_0^{2\pi} K(t,s)(\tfrac{1}{2}\sin 2t)\,dt = \tfrac{1}{2}\sin 2t = \psi_2(s);$$

$$(1/\pi)\int_0^{2\pi} K(s,t)(\sin t + 2\sin 2t)\,dt = 2\sin s + (\sin s + 2\sin 2s) = \varphi_1(s) + \varphi_2(s);$$

$$(1/\pi)\int_0^{2\pi} K(t,s)(\tfrac{1}{2}\sin t + \tfrac{1}{4}\sin 2t)\,dt = (\tfrac{1}{2}\sin s + \tfrac{1}{4}\sin 2s) + \tfrac{1}{2}\sin 2s = \psi_1(s) + \psi_2(s).$$

Now let $G(s,t)$ be an arbitrary function represented by the series $\sum_1^\infty c_i \sin is + \sum_0^\infty d_i \cos is$, convergent in the mean. Then $K(s,t)$ is of positive type with respect to $G(s,t)$. For

$$\int_0^{2\pi}\int_0^{2\pi} [(2\sin s)(\tfrac{1}{2}\sin t + \tfrac{1}{4}\sin 2t) + (\sin s + 2\sin 2s)(\tfrac{1}{2}\sin 2t)] \times [\sum c_i \sin is + \sum d_i \cos is]\,[\sum c_i \sin it + \sum d_i \cos it]\,ds\,dt$$

$$= \pi\int_0^{2\pi} [2c_1(\tfrac{1}{2}\sin t + \tfrac{1}{4}\sin 2t) + (c_1 + 2c_2)\tfrac{1}{2}\sin 2t] \times (\sum c_i \sin it + \sum d_i \cos it)\,dt$$

$$= \pi^2(c_1^2 + c_1 c_2 + c_2^2)$$

and the quadratic form in the right member is of positive type.

The function $K(s,t)$ of this example is of course not positive *definite* since it is orthogonal to $\cos is$ for every i and to $\sin is$ for $i \geqq 3$. But the function

$$K(s,t) + \sum_{i=3}^{\infty} a_i \sin is \sin it + \sum_{i=0}^{\infty} b_i \cos is \cos it,$$

where the a's and b's are positive constants, is clearly positive definite, and we have an example of a function that is non-symmetric and yet positive definite.

Under the hypotheses of Theorem III, each principal kernel $k_i(s,t)$ reduces to the sum of the pairs of fundamental functions corresponding to the characteristic constant λ_i, and the series $\sum k_i(s,t)$ coincides with the series proved convergent in Theorem III of our earlier paper* (under the assumption that all the fundamental functions have the squares of their integrals less than some number N independent of i). Then†

$$K(s,t) = \sum k_i(s,t) + H(s,t),$$

* Loc. cit., p. 175.

† Goursat, loc. cit., p. 428. Our series is convergent in the mean, whereas Goursat's is uniformly convergent, but this does not effect the present results.

where $H(s, t)$ is orthogonal to $\sum k_i(s, t)$. If the sets of orthogonal functions Φ_i and Ψ_i are closed, $H(s, t)$ is zero (except perhaps on a set of measure zero). Suppose moreover, that there is only one fundamental function corresponding to each λ_i, that is, that each is a root of $D(\lambda)$ of multiplicity 1. Then the fundamental functions Φ_i and Ψ_i (normalized) coincide with the fundamental functions of Schmidt. For*

$$\begin{aligned}\lambda_j \int_a^b K(s, \tau)\, \Psi_j(\tau)\, d\tau &= \lambda_j \int_a^b \sum \{(1/\lambda_i)\, \Phi_i(s)\, \Psi_i(\tau)\}\, \Psi_j(\tau)\, d\tau, \\ &= \Phi_j(s), \\ \lambda_j \int_a^b K(\tau, s)\, \Phi_j(\tau)\, d\tau &= \lambda_j \int_a^b \sum \{(1/\lambda_i)\, \Phi_i(\tau)\, \Psi_i(s)\}\, \Phi_j(\tau)\, d\tau \\ &= \Psi_j(s).\end{aligned}$$

* Goursat, loc. cit., p. 470.

THE TSCHIRNHAUS TRANSFORMATION.*

By Raymond Garver.

1. **Introduction.** It is the purpose of this article to collect results which have been obtained by applying the Tschirnhaus transformation

$$y = k_1 x^{n-1} + k_2 x^{n-2} + \cdots + k_{n-1} x + k_n \tag{1}$$

to the nth degree equation

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0. \tag{2}$$

Some mention is also made of rational transformations; this is justified by the well-known fact that a rational transformation is equivalent to a Tschirnhaus transformation. (See any of the following references in the bibliography: 2, 22, 34, 38, 78, 84, 96.)

Equation (1) is to be interpreted as a transformation on the roots of (2); each root x_i $(i = 1, 2, \cdots, n)$ gives a value y_i, and these y_i are the roots of a new nth degree equation in y, called the transformed equation. The coefficients of the transformed equation involve the k_i of (1), which we regard as parameters. It may then be possible to so choose the k_i that the new equation in y will take a particularly simple form or one that possesses certain special properties.† Herein lies the value of the Tschirnhaus transformation. The first writers on the subject limited themselves to the problem of making as many coefficients as possible of the transformed equation vanish, but later other results were sought.

The problem of setting up the transformed equation is one that I shall not consider at length; it is treated rather completely in references 22, 83, 84 and 92, and in less detail, or for particular cases, in references 27, 38, 46, 48, 72, 73, 92, 96 and 98. We may, however, by outlining one method learn something about the coefficients with no actual computation. This method depends on the fact that the coefficients of the new equation will be known when we know $\sum y_i$, $\sum y_i^2$, $\sum y_i^3$, etc. And these can be found directly from (1) and (2). For example, if we use s_j to represent Σx_j, we have at once

$$\sum y_i = k_1 s_{n-1} + k_2 s_{n-2} + \cdots + k_{n-1} s + n k_n. \tag{3}$$

* Received October 18, 1927.

† The choice of the k_i will depend on the solution of a number of equations of different degree in the k_i, and the question of the degree of the various equations which must be solved before the k_i can be completely determined is a very important one.

To find $\sum y^2$ (henceforth I shall omit the subscript i in a symbol of this kind) we square both sides of (1) and sum over the roots of (2), and so on. It is then clear that the coefficient of y^{n-1} in the transformed equation is of the first degree in the k_i, that of y^{n-2} is of the second degree, etc.

Another incidental problem, which is of importance if we are using the Tschirnhaus transformation to solve an equation of type (2), is the determination of the relation between the roots of (2) and those of the transformed equation. Briefly, we may say that if the transformed equation has distinct roots, the roots of the given equation may be expressed rationally in terms of them, while if the transformed equation has a root of multiplicity m, the expression of the roots of the given equation in terms of those of the transformed equation will involve the solution of an equation of degree m. This point is covered, in more or less detail, in references 34, 46, 73, 78, 83, 84, 92 and 96.

2. **The work of Tschirnhaus.** The idea of removing more than one term from an equation by means of a suitably chosen transformation seems to have originated with Tschirnhaus.* His published article on the subject appeared in 1683 (ref. 94), but six years before this he had outlined his method in a letter to Leibniz (ref. 74, page 333). He wrote, in part, "Si jam velimus duos terminos in quacunque aequatione auferre, supponendum $x^2 = ax+b+y$, si tres, $x^3 = ax^2+bx+c+y$, si quatuor $x^4 = ax^3+bx^2+cx+d+y$,† atque sic in infinitum." He then shows how his method leads to the solution of the general cubic equation. However, he believed, or professed to believe,‡ that his method led to the algebraic solution of equations of any degree, that is, that the elimination of the intermediate coefficients would not require the solution of any equations of degree greater than $n-1$. In a later letter to Leibniz (ref. 74, page 359) he mentioned it as one of three methods which he had devised for the solution of equations, and particularly quintics.

Leibniz perceived that the method would not be successful, and replied, in an undated letter which was apparently sent in 1679 or 1680 (ref. 74, page 403), "Tertiam tuam methodum radices aequationem inveniendi, nempe si sit $x^5+px^4+qx^3+rx^2+sx+t = 0$, ponendo $x^4+bx^3+cx^2+dx+e = y$, tollendo x per y et ope arbitrarium b, c, d, etc. auferendo terminos intermedios in aequatione proveniente $y^5+\cdots$ etc. $= 0$, non puto

* Vieta and Cardan had previously showed how to remove one term by a linear transformation.

† I have used modern notation in writing these equations, as well as those in the letter of Leibniz below. Our $=$ sign was not then in use, and Tschirnhaus used xx instead of x^2.

‡ Cantor (ref. 23) has an interesting discussion of this point. See also Zeuthen, ref. 100, pages 65 and 233.

succedere posse in altioribus nisi quoad casus speciales. ··· Utile esse potest ad aequationes transformandas, non tamen (generaliter) ad resolvandas."

However, Tschirnhaus published his article, with the claim that it allowed the roots of equations of any degree to be determined.* He repeated this claim, with regard to the quintic, in a later letter to Leibniz (ref. 74, pages 449 and 451). But all that he actually accomplished was to reduce the general cubic to binomial form, and to give a general formula which will remove the second and third coefficients from an equation of any degree. This transformation requires the solution of a quadratic equation only, and may be found in references 23, 74 and 98. Such a reduced equation we shall call, after Klein, a principal equation. Tschirnhaus did not mention specifically any particular equations of degree greater than 3.

3. **Transformations on cubic equations.** In our notation, the solution of the cubic equation by the method of Tschirnhaus depends on three steps: 1) setting up and solving the equations $\sum y = \sum y^2 = 0$, by proper choice of k_2 and k_3 (k_1 may be taken equal to 1), 2) computing A_3 for the transformed equation $y^3 + A_3 = 0$, 3) determining the relation between the roots of the original equation and those of the new. The solution is given completely by Lagrange (ref. 72), Weber (ref. 96), and myself (ref. 42). It is also discussed, in more or less detail, in many other articles.†

Among other results which have been obtained in connection with cubic equations, one of the most important is due to Hermite. It is a special case of a general theorem which we shall have occasion to refer to later. For the cubic, the theorem shows that, if we write the equation in the form $a_0x^3 + 3a_1x^2 + 3a_2x + a_3 = 0$ and put $y = t_1(a_0x + a_1) + t_0(a_0x^2 + 3a_1x + 2a_2)$, the transformed equation will have the form $y^3 + A_2y + A_3 = 0$, where A_2 and A_3 are covariants of the second and third order respectively of the form $a_0t_1^3 + 3a_1t_1^2t_0 + 3a_2t_1t_0^2 + a_3t_0^3$. Since a cubic form has but one covariant of each order, the coefficients A_2 and A_3 are at once determined except for a constant factor, which can be found easily by taking a special case. The details of the Hermite transformation are given by Cayley, Fricke and Weber, and briefly by Salmon and Faà di Bruno. (See references 25, 27, 38, 96, 82 and 37.)

Tortolini (ref. 93) has employed a quadratic transformation to reduce the general cubic to the solvable form $(x + p)^3 = p^3 - r$. He takes $k_1 = 1$,

* Weissenborn (ref. 98, page 25) gives a reference from the Tschirnhaus article which substantiates this statement.

† See references 23, 24, 46, 59, 76, 93 and 98. Briefer treatments will be found in 2, 22, 48, 77, 80, 83, 84.

$k_3 = 0$, after which k_2 must be chosen to satisfy a quadratic equation. Serret (ref. 84), Simonart (ref. 86) and Niewenglowski (ref. 78) have employed a linear fractional transformation to reduce the cubic to binomial form, and hence to solve it. Niewenglowski and Burnside and Panton (ref. 22) state that the general Tschirnhaus transformation relative to the cubic equation is equivalent to a linear fractional transformation. There are, however, a few exceptional Tschirnhaus transformations for which this is not the case. Cayley (ref. 31) has discussed transformations of a cubic into itself; he thus finds a field in which any cubic can be considered as an Abelian equation. I have considered (ref. 42) transformations on a number of special cubics.

4. Transformations on quartic equations. The first application of the Tschirnhaus transformation to quartic equations we made by Lagrange (ref. 72), almost a hundred years after Tschirnhaus. He first showed that the quartic may be solved by this method. For, applying a quadratic transformation $y = x^2 + k_3 x + k_4$, we can choose k_3 and k_4 (by solving a linear and a cubic equation) so that $\sum y = \sum y^3 = 0$. The transformed equation is then of the solvable form $y^4 + A_2 y^2 + A_4 = 0$. Lagrange gives the details of the reduction, as do Bring (see below), Hirsch (ref. 59) and Weissenborn (ref. 98). Briefer mention is found in a number of references.*

Lagrange also considered the reduction of the quartic to binomial form by means of a cubic transformation $y = x^3 + k_2 x^2 + k_3 x + k_4$. The equations of condition, $\sum y = \sum y^2 = \sum y^3 = 0$, being of degree 1, 2 and 3 in the k_i respectively, we should expect to arrive finally at an equation of the 6th degree in one of them, say k_2. Lagrange gives an elegant a priori proof that this sextic will decompose into three quadratic factors, of which the coefficients are themselves the roots of cubic equations. Hence the reduction can be effected without requiring the solution of equations of degree greater than three. This decomposition is undoubtedly a special result for the quartic; it would not follow for higher degree equations. And even for the quartic it would be very laborious to actually obtain and factor the required sextic. I have exhibited a method (ref. 40) whereby the actual reduction to binomial form might be more easily carried through.

The Swedish mathematician and historian Bring, whose chief work was in connection with the quintic, also considered quartic equations (ref. 5). He first used a quadratic transformation to obtain the form $y^4 + A_3 y + A_4 = 0$, and then reduced this to Lagrange's solvable form. His work, however, was undoubtedly independent of that of Lagrange, for he also attempted

* Refs. 2, 22, 37, 46, 48, 76, 77, 80, 84, 96.

to reduce the quartic to binomial form, but was unsuccessful. He did apply a reciprocal transformation to his form above, and thus obtained a quartic lacking its third and fourth coefficients.

Hermite's modified form of the Tschirnhaus transformation (see above under cubics, also ref. 53) has also been studied in connection with quartic equations. If we take the equation in the form $(a_0, a_1, a_2, a_3, a_4)(x, 1)^4 = 0$,* and apply the transformation $y = t_2(a_0x + a_1) + t_1(a_0x^3 + 4a_1x + 3a_2) + t_0(a_0x^3 + 4a_1x^2 + 6a_2x + 3a_3)$, the second term of the transformed equation will be zero, while the other coefficients will be simultaneous invariants of the two forms $(a_0, a_1, a_2, a_3, a_4)(x, y)^4$ and $(t_0, -t_1, t_2)(x, y)^2$. Cayley has computed the transformed quartic (refs. 26 and 27), and he later determined the second and third degree invariants of the transformed quartic (ref. 28). He was then able to show that the effect of the Tschirnhaus transformation on any quartic $U = 0$ is to reduce it to a quartic which is equivalent under a linear fractional transformation to a linear combination of U and its Hessian. Most of Cayley's results are given by Salmon, and some by Burnside and Panton.

Hermite himself used a transformation slightly different from the one indicated above to reduce the quartic to the form $x^4 - 6Sx^2 - 8Tx - 3S^2 = 0$. (See ref. 53.) He then solved this equation by means of elliptic functions. Tortolini (ref. 93) later discussed the same normal form.

Wiman (ref. 99) has showed that the normal form $y^4 + A_2y^2 + A_4 = 0$ may be obtained by a rational transformation (and hence the coefficients A_2 and A_4 will be rational if the original quartic has rational coefficients) provided the quartic has rational coefficients, is irreducible in the field of rational numbers, and has an imprimitive Galois group for that field. I have given a different proof of the same result, employing a quadratic transformation only (ref. 41). Bucht (ref. 21) has pointed out further that if the Galois group is the so-called 4-group A_4 will be not only a rational number, but also the square of a rational number.

Burnside and Panton give a number of exercises in connection with Tschirnhaus transformations. One of them states that the general Tschirnhaus transformation of a quartic equation is equivalent to a rational transformation of the form $y = P/(x - a) + Q/(x - b)$. Sylvester (ref. 87) has given a more general theorem concerning the equivalence of Tschirnhaus transformations and fractional transformations.

5. **Transformations on quintic equations.** Passing now to the quintic, we find a number of normal forms, several of which have been used in solving the quintic by non-algebraic methods. The principal equation

* This notation means $a_0x^4 + 4a_1x^3 + 6a_2x^2 + 4a_3x + a_4 = 0$.

of the fifth degree, which can be obtained by solution of a single quadratic, may be credited to Tschirnhaus. Several solutions of the quintic involve only a reduction to the principal form, for example, those of Klein (ref. 69), Gordan (ref. 43) and Heyman (ref. 56–57). The Bring-Jerrard normal form $y^5+A_4y+A_5=0$, which requires the solution of a cubic, was first obtained by Bring in 1786, but his work was forgotten. Jerrard (ref. 61) gave a different reduction in 1834, and the interest in his work led to a rediscovery of that of Bring.* If we apply $y=k_1x^4+k_2x^3+k_3x^2+k_4x+k_5$ to the general quintic equation and attempt to make $\sum y=\sum y^2=\sum y^3=0$ we shall arrive at an equation of the sixth degree. Bring avoids this by first reducing the quintic to principal form, choosiug $k_1=1$, choosing k_5 so that $\sum y=0$, and showing that by proper choice of k_3 and k_4 as linear functions of k_2 the equation $\sum y^2=0$ vanishes identically in k_2. The third condition $\sum y^3=0$ is then a cubic in the single remaining parameter k_2. Jarrard's process, after choosing k_5 as above, is somewhat different. He lets k_1 remain, so that $\sum y^2$ is a homogeneous, quadratic function of k_1, k_2, k_3 and k_4. By a known theorem† it can then be expressed in the form $P^2+Q^2+R^2+S^2$, where P, Q, R and S are linear (homogeneous) functions of the four k_i. The condition $\sum y^2=0$ may be satisfied by taking $P=iQ$, $R=iS$, which will thus determine two of the k_i linearly in terms of the other two. And $\sum y^3=0$ then becomes a cubic in the ratio of the two remaining k_i. By solving the quartic $\sum y^4=0$ instead of the cubic $\sum y^3=0$ Jerrard showed how to obtain the normal form $y^5+A_3y^2+A_5=0$, and then by means of reciprocal transformations he arrived at two additional forms. Any of these trinomial forms may be changed into a form with a single parameter by putting $y=ct$, and choosing c properly.

Jerrard's work aroused considerable interest; Hermite said, in using the Jerrard form $x^5-x-a=0$ to indicate the first solution of the quintic by means of elliptic functions (see ref. 52), "Ce résultat remarquable (i. e., the reduction of Jerrard) est le pas le plus important qui ait été fait dans la théorie algébrique des équations du cinquième degré, depuis qu'Abel a démontré qu'il était impossible de les résoudre par radicaux." Unfortunately Jerrard somewhat dimmed his fame by publishing a number of false results concerning supposed algebraic solutions of the quintic. Some of these involved Tschirnhaus transformations; thus he believed that the general

* Grunert (ref. 47) and Harley (ref. 50) may be consulted in this connection. For other comments on Bring, see Klein (ref. 69, page 157), Sylvester (ref. 87, page 531 and ref. 89 page 553), and reference 23.

† For proofs or general statements of this theorem see ref. 22, page 174; 46, 227; 76, 123; 79, 84; 83, 82; 84, 430; 92, 253; 95, 162.

quintic could be reduced, by such a transformation, to the solvable De Moivre form $x^5+Bx^3+B^2x/5+E=0$ (see refs. 62 and 63).

The Bring-Jerrard quintic is discussed in a large number of references. Many of them do no more than reproduce the work of Jerrard, in more or less detail.* Heal (ref. 51) and Lavagna (ref. 73) give calculations in connection with the Jerrard transformation which are not found elsewhere; the latter calculated the coefficients of the transformed equation. Sievers (ref. 85) gives a reduction similar to that of Bring; he assumes k_3 a certain linear combination of k_2 and k_4, after which k_4 is determined as the root of a quadratic equation and k_2 as the root of a cubic. Cayley (ref. 27) obtains the desired transformation in a very different way. A number of writers have interpreted the Tschirnhaus transformation geometrically, and have been able to derive the Bring-Jerrard quintic by a geometrical argument. Among these are Cayley (ref. 27, page 391), Dickson, Clebsch (ref. 32, page 328), Klein (ref. 69, page 230–1) and Weber (ref. 96, page 204–6). Matthiessen (ref. 76, page 125) and Salmon (ref. 82, page 250) make a statement which is almost certainly incorrect, namely that the Bring-Jerrard form can be obtained by a cubic transformation, without solving any equation higher than a cubic. It is also necessary to mention Hamilton's inquiry into the validity of Jerrard's work (see ref. 49). Among a great many other results, he pointed out that the Jerrard method is not valid when the given equation is a quartic, and that the coefficients of the transformed equation will often be imaginary, even when those of the original equation are real. This latter point has also been considered by Catalan (ref. 24) and Sylvester (refs. 87 and 89), who showed that the new coefficients will be imaginary unless the given quintic (with real coefficients) has four imaginary roots.

There remain to discuss three other interesting transformations which lead to the Bring-Jerrard quintic. One of these is of the form $y=u\varphi+v\psi$, where φ and ψ are polynomial functions of x which possess the five properties $\sum\varphi=\sum\psi=\sum\varphi^2=\sum\psi^2=\sum\varphi\psi=0$. (Such functions can be shown to exist.) We then have $\sum y=\sum y^2=0$ at once, and $\sum y^3=0$ becomes a cubic in the ratio u/v. This transformation was first used by Gordan, in 1885, (see ref. 44), and similar, though by no means identical, reductions have been given by Dickson, and myself (ref. 39). Transformations of the same type have also been studied by Brioschi (ref. 12).

The general quintic can not be transformed to the Bring-Jerrard form by means of a linear fractional transformation; a necessary and sufficient condition that it can is that one of its twelfth degree invariants vanishes

* See refs. 2, 22, 23, 24, 46, 48, 76, 77, 83, 84, 92, 95, 98.

(see Elliott, ref. 35). Based on this fact is a general reduction outlined by Clebsch (ref. 32), which employs a preliminary Tschirnhaus transformation to obtain a quintic for which this invariant does vanish.

Finally we consider a transformation due to Hermite (ref. 55). He first proves a general theorem: If $f(x, y)$ is a form of the nth order, and $\varphi(x, y)$ one of its covariants of order $(n-2)$,* and if we apply the transformation $z = \varphi(x, 1)/f'_x(x, 1)$ to the equation $f(x, 1) = 0$, the coefficients of the transformed equation in z will all be invariants of $f(x, y)$.† Brioschi has, in a number of articles (refs. 9, 10, 11, 13, 15, 18), considered transformations of the type mentioned in the theorem, both on quintics and sextics. For the reduction to the Bring-Jerrard form, however, Hermite required a more elaborate form. He thus took

$$z = \frac{t\varphi_1(x, 1) + u\varphi_2(x, 1) + v\varphi_3(x, 1) + w\varphi_4(x, 1)}{f'_x(x, 1)},$$

where t, u, v and w are parameters, and the φ_i are covariant cubics of respective degrees 3, 5, 7 and 9. A large number of terms in the transformed equation vanish (this can be deduced from the invariantive properties of its coefficients), and the parameters can be so chosen that the Bring-Jerrard form is obtained. Cayley (ref. 29) and Salmon give excellent summaries of Hermite's work. Rahts (ref. 81) continued Hermite's work, correcting a few numerical mistakes, and carrying out many additional calculations. He computed explicitly the transformed equation.

I have already mentioned another theorem of Hermite, in connection with the cubic and quartic. It hardly seems necessary to state the theorem in general; it can almost be deduced from the statement given in section 4 on quartics. Or the reader may consult Faà di Bruno (page 269 ff.), Fricke (page 167 ff.) or Weber (page 240 ff.). Weber applies this transformation of Hermite, with slight modifications, to the quintic, and shows that by means of it the Bring-Jerrard form may be obtained (page 260 ff.). Weber and Cayley (ref. 27) seem to be the only writers to apply this type of transformation to equations of higher than the fourth degree.

The other important quintic normal form is the Brioschi form, which we may write as $y^5 + 10By^3 + 45B^2y + k = 0$. The name is, in a sense, inaccurate; it would be more correct to say that the Brioschi resolvent $y^5 + 10By^3 + 45B^2y + k = 0$ can be shown to be a quintic normal form. For Brioschi considered the equation not in this connection, but as a resolvent

* Such covariants exist when n is at least 5.

† Faà di Bruno has given a generalization of this theorem (page 191). See also Brioschi (ref. 11, page 293).

for certain Jacobian sextic equations. His first article on the subject (ref. 6), published in 1858, contains numerical inaccuracies; these are corrected in reference 9. The correct resolvent had also been given previous to the publication of this latter article by Hermite (ref. 55), and Joubert (ref. 64). It seems customary to credit the rectification of the coefficients to Joubert (see, for example, Klein, ref. 69, page 250; Hagen, page 342), though Hermite's results were published slightly earlier.

The Brioschi normal form may be said to be characterised by two properties; the second and fourth coefficients are missing, and it may be obtained by means of a transformation involving only square roots. The first forms of this general type were obtained by Kronecker (refs. 70 and 71) and Hermite (ref. 54), in 1861. The transformed equation of Hermite was $y^5 + T_{12}y^3 + T_{16}\Delta y + T_{18}\sqrt{\Delta^3} = 0$, in which the T_i are invariants of the given quintic, the degree being indicated by the subscript, and Δ is the discriminant of the given quintic. The transformation which he employed was not a Tschirnhaus transformation, but Brioschi later showed (see ref. 9, page 144 and ref. 10, page 305) that it was equivalent to a Tschirnhaus transformation, or more exactly to a transformation of the type $\varphi(x, 1)/f'_x(x, 1)$ mentioned above. Brioschi also used transformations of this same type in a number of articles, as has been mentioned above. None of the quintic forms that he obtains are particularly simple.

Hermite was also the first to actually obtain the Brioschi quintic as a quintic normal form (see ref. 55). However, he did not follow the method which he used in obtaining the form indicated just above, but rather, as he himself points out, based his work on that of Kronecker. His method of setting up the necessary transformation is long and roundabout; in fact, even the simplest reduction is so long that I shall not attempt to give any of the details. Other derivations of the Brioschi form were soon given by Kiepert (refs. 65 and 66), Gordan (refs. 43, 44 and 45), Klein (ref. 69, page 170), whose method follows that of Hermite, and Heyman (ref. 56). Of these, the most direct reduction is that given by Gordan in 1887 (ref. 45). Weber (ref. 96, page 263 ff.) later succeeded in simplifying Gordan's work somewhat, reducing the number of irrationalities used by one. Dickson has simplified Weber's method slightly, and his presentation is undoubtedly the most satisfactory available. It includes a precise determination of the irrationalities involved.

The Brioschi quintic is perhaps more conveniently solved by elliptic functions than is the Bring-Jerrard form. The details of the solution are given in concise form by Kiepert (refs. 65 and 66).

In conclusion I mention the use of the Tschirnhaus transformation by Klein (ref. 69, pages 171–2 and 286–7) in his proof of a theorem of

Kronecker, namely that it is impossible, in the case of the general quintic, "even after adjunction of the square root of the discriminant, to construct a rational resolvent which contains only one parameter".

6. **Transformations on higher degree equations.** When we come to equations of higher degree, we find investigations of two types, those on equations of some particular degree, and those of a more general nature. Brioschi has, in several articles (see refs. 11, 14, 16, 17 and 19), considered "de quelle manière on peut transformer une équation générale du sixième degré au double point de vue d'annuler un ou plusieurs des coefficients de la transformée et de rendre les autres des invariants de l'équation donnée".* Brill (ref. 4) and Maschke (ref. 75) have also considered sextic normal forms; a summary of Brill's work is given by Salmon (ref. 82, page 276 ff.). Transformations on Jacobian sextic equations have been considered by Kronecker (ref. 70), in connetion with his solution of the quintic, Brioschi (refs. 8 and 9), Kiepert (ref. 66) and Cayley (ref. 30).† Brioschi discusses transformations on seventh degree equations in references 12 and 20. Hilbert (ref. 58) has a result on ninth degree equations, which we shall have occasion to refer to later.

There remain for discussion two investigations of a more general nature. The work of Lagrange (ref. 72) has already been mentioned in connection with the cubic and quartic equation, but his article goes much farther. Writing before the algebraic solution of the quintic and higher degree equations had been proved to be impossible, he was still interested (as Tschirnhaus had been) in the method as one which might possibly lead to such a solution. He did, however, admit that he had little hope of success.‡ His investigations are none the less of interest. In order to solve the general equation of any prime degree n, he considered the reduction to binomial form. If this could be completed without requiring the solution of any equation of degree higher than $n-1$, the solution of the given equation could be carried through in a manner similar to that which we have already indicated for the cubic. But this is not the case; if we attempt to remove the four intermediate terms of a quintic we arrive at an equation of degree $1\times2\times3\times4 = 24$, and in general to remove the intermediate terms from an nth degree equation we shall arrive at an equation of degree $(n-1)!$. Lagrange was able to show, by a method similar to that which he used in the reduction of the quartic to binomial form, that this equation will factor into $(n-2)!$ equations of degree $n-1$, but the coefficients of these new equations are themselves roots of equations

* Ref. 17, page 320.

† See also Klein (ref. 69, pages 167–8), and Hagen (ref. 48, pages 340–2).

‡ See page 306.

of degree $(n-2)!$. Thus for the quintic we still have equations of the sixth degree to deal with. When n is composite, it is unnecessary to remove all intermediate terms; the general results, though, for this case are difficult to write down, since there are many different possibilities. The results for the simplest case, the sextic, are easily given. The general sextic could be solved if we could obtain either of the forms $y^6+A_2y^4+A_4y^2+A_6=0$ or $y^6+A_3y^3+A_6=0$ without encountering high degree equations. To reach the first will clearly require the setting up of an equation of degree $1\times3\times5=15$; for this case Lagrange indicates no reduction of degree. An attempt to obtain the second form will lead to an equation of degree $1\times2\times4\times5=40$, but Lagrange shows that this equation will factor into quadratics whose coefficients are themselves roots of equations of the 20th degree.

Sylvester (refs. 87, 88 and 89) considers a problem of a different nature, one which Hamilton (ref. 49) had already treated for certain special cases. He wants to know how high degree an equation we must have before k successive terms after the first can be removed by means of a Tschirnhaus transformation without requiring the solution of any equation of degree greater than k. His results may not be absolutely correct, for he is actually considering a somewhat different problem. Thus he states that 5 is the minimum degree of an equation which admits of 3 terms being removed without solving an equation of above the 3rd degree. And we have already seen that a similar reduction is possible for the quartic. The difference in results lies in the fact that Sylvester does not take into consideration the possible factorization of a sextic equation. He is instead considering the determination of the number of parameters which must be present before a general system of three equations of degree 1, 2 and 3 can be solved without introducing any elevation of degree. The problems are equivalent except for the possibility that, in a special case, an elevation of degree may be counteracted by a subsequent factorization. For the removal of 4 terms Sylvester found that an equation of the 11th degree is required if we are to allow the introduction of a single quartic equation (this he calls a solution of "minimum weight"), while a 10th degree equation will suffice if we allow a second quartic to enter. This result had been given by Hamilton. Hilbert (ref. 58) has indicated a method by which four terms may be removed from a 9th degree equation, but this requires the solution of a quintic. For the removal of more than 4 terms, there are again two results, depending on whether the system of auxiliary equations is of minimum weight or not. The numbers corresponding to the minimum weight solutions are called by Sylvester "Hamilton's numbers"; he and Hammond compute 10 of them, while Hamilton had already obtained the

first six. They run 2, 3, 5, 11, 47, 923, 409619, 83763206255, etc. The next number has 22 digits, and the tenth, which represents the minimum degree of equation from which ten successive terms after the first can be removed without requiring the solution of more than one 10th degree equation, has 43. In each case, after the third, a reduction is possible if we allow more of the higher degree equations to enter: the reduced numbers are 2, 3, 5, 10, 44, 905, 409181, 83762797734, etc.

Bibliography.

(Books and articles which have not been available during the preparation of this article are marked with a *).

1. Ball, Notes on Biquadratic Equations, Quar. Jour., 7, 1866, 358–369.
2. Bauer, Vorlesungen über Algebra, Leipzig, 1903, chapter XI.
3. *Blomstrand, De praecipuis methodis, Lundae, 1847.
4. Brill, Über binäre Formen und die Gleichung sechsten Grades, Math. Annalen, 20, 1882, 330–356.
5. Bring, Meletemata ..... circa transformationem aequationem, Lund, 1786. The important parts of Bring's work are reproduced in Grunert's Archiv, 41, 1864, 105–112, and in Annali di Mat., ser. 1, 6, 1864, 33–42. His work on quintics is also given in Quar. Jour., 6, 1864, 45–47.
6. Brioschi, Sulla risolucione delle equazioni del quinto grado, Annali di Mat., ser. 1, 1, 1858, 256–259 and 326–328. Also in his Œuvres, vol. I, Milan, 1901, 335–341.
7. Brioschi, Sulla trasformazione delle equazioni algebriche, Atti del R. Istituto Lombardo, 1, 1858, 231–233; Œuvres, vol. III, 1904, 171–175.
8. Brioschi, La soluzione più generale delle equazioni del quinto grado, Annali de Mat., ser. 2, 1, 1867–8, 222–231; Œuvres, vol. II, 1902, 79–89.
9. Brioschi, Über die Auflösungen der Gleichungen vom fünften Grade, Math. Annalen, 13, 1878, 109–160.
10. Brioschi, Appendice terza to his translation of Cayley's Elliptic Functions, Milan, 1880; Œuvres, vol. IV, 1906, 260–322.
11. Brioschi, Sulle relazioni esistenti fra covarianti ed invarianti di una stessa forma binaria, Annali di Mat., ser. 2, 11, 1882–3, 291–304; Œuvres, vol. II, 281–294.
12. Brioschi, Sulla trasformatione delle equazioni algebriche, Rendiconti del R. Istituto Lombardo, ser. 2, 20, 1887, 364–370; Œuvres, vol. III, 293–299.
13. Brioschi, Über die Transformation der Algebraischen Gleichungen durch Covarianten, Math. Annalen, 29, 1887, 327–330; Œuvres, vol. V, 1909, 249–252.
14. Brioschi, La forma normale delle equazioni del sesto grado, Rendiconti della R. Accademia dei Lincei, ser. 4, 4, 1888, 301–305 and 485–488; Œuvres, vol. IV, 43–52.
15. Brioschi, Sopra una trasformazione delle equazioni del quinto grado, Annali di Mat., ser. 2, 16, 1888–9, 181–189; Œuvres, vol. III, 21–30.
16. Brioschi, Principii di una teoria sulla trasformazione delle equazioni algebriche, Annali di Mat., ser. 2, 16, 1888–9, 329–334; Œuvres, vol. III, 31–37.
17. Brioschi, Sur l'équation du sixième degré, Acta Mathematica, 12, 1889, 83–101; Œuvres, vol. V, 313–330.
18. Brioschi, Sur la transformation des équations algébriques, Proceedings of London Math. Soc., 20, 1889, 127–131; Œuvres, vol. V, 361–365.

19. Brioschi, Sur l'équation du sixième degré, Annales Scientifiques de l'École Normale Supérieure, ser. 3, 12, 1895, 343–350; Œuvres, vol. V, 175–181.
20. Brioschi, Sur la transformation des équations algébriques, Comptes Rendus, 124, 1897, 661–665; Œuvres, vol. V, 75–80.
21. Bucht, Über einige algebraische Körper achten Grades, Arkiv för Matematik, Astronomi, och Fysik, 6, 1910–11, No. 30.
22. Burnside and Panton, Theory of Equations, 4th ed., vol. II, Dublin, 1901, chapter XIX.
23. Cantor, Geschichte der Mathematik, vol. III, Leipzig, 1898, and vol. IV, 1908. Page references are found in the index.
24. Catalan, Sur la transformation des équations, Nouvelle Correspondance Mathématique, 2, 1876, 342–349.
25. Cayley, Note sur la transformation de Tschirnhausen, Crelle's Journal, 58, 1861, 259–262. Also in his Collected Math. Papers, vol. IV, Cambridge, 1891, 364–367.
26. Cayley, Deuxième note sur la transformation de Tschirnhausen, Crelle's Journal, 58, 1861, 263–269; Papers, vol. IV, 368–374.
27. Cayley, On Tschirnhausen's Transformation, Phil. Trans., 151, 1861, 561–578; Papers, vol. IV, 375–394.
28. Cayley, Addition to the Memoir on Tschirnhausen's Transformation, Phil. Trans., 156, 1866, 97–100; Papers, vol. V, 1892, 449–453.
29. Cayley, An Eighth Memoir on Quantics, Phil. Trans., 157, 1867; Papers, vol. VI, 1893, 147–190. Especially articles 286–292.
30. Cayley, On the Jacobian Sextic Equation, Quar. Jour., 18, 1882, 52–65; Papers, vol. XI, 1896, 389–401.
31. Cayley, Note on a Cubic Equation, Messenger of Math., 15, 1886, 62–64; Papers, vol. XII, 1897, 421–423.
32. Clebsch, Über die Anwendung der quadratischen Substitution auf die Gleichungen 5ten Grades, Math. Annalen, 4, 1871, 284–345.
33. Cockle, On the Transformation of Algebraic Equations, Mathematician, July, 1844.
34. Dickson, Modern Algebraic Theories, Chicago, 1926, chapter XII.
35. Elliott, Algebra of Quantics, 2nd ed., Oxford, 1913. Especially pages 261–262.
36. Encyklopädie der Mathematischen Wissenschaften, vol. I, part I, Leipzig, 1898–1904. Page references are found in the index to vol. I, which is at the end of part II.
37. Faà di Bruno, Theorie der binären Formen, German translation by Walter, Leipzig, 1881. Especially page 191 and page 269 ff.
38. Fricke, Lehrbuch der Algebra, vol. I, Braunschweig, 1924, chapter VI.
39. Garver, Transformations of one Principal Equation into another, Annals of Mathematics, ser. 2, 28, 1927, 112–116.
40. Garver, The Binomial Quartic as a Normal Form, Bulletin of the American Mathematical Society, 33, 1927, 677–680.
41. Garver, A Rational Normal Form for certain Quartics, Bulletin of the American Mathematical Society, 34, 1928, 73–74.
42. Garver, Tschirnhaus Transformations on certain rational Cubics, American Mathematical Monthly, 34, 1927, 521–525.
43. Gordan, Über die Auflösung der Gleichungen vom fünften Grade, Math. Annalen, 13, 1878, 375–404.
44. Gordan, Sur les équations du cinquième degré, Journal de Math., ser. 4, 1, 1885, 455–458.
45. Gordan, Über Gleichungen fünften Grades, Math. Annalen, 28, 1887, 152–166.
46. Grunert, Die Methoden von Tschirnhaus und Jerrard zur Transformation der Gleichungen, Grunert's Archiv, 40, 1863, 214–231.

47. Grunert, Wichtige historische Mitteilung, Grunert's Archiv, 40, 1863, 515–516. See also reference 5 of this bibliography.
48. Hagen, Synopsis der Höheren Mathematik, vol. I, Berlin, 1891. Page references are found in the index.
49. *Hamilton, Sixth Report of the British Association for the Advancement of Science, London, 1837, 295–348.
50. Harley, Reduction of the General Equation of the Fifth Degree to a Trinomial Form, Quar. Jour., 6, 1864, 38–45.
51. Heal, On the Removal of Terms from an Equation of the Fifth Degree, The Analyst, 6, 1879, 78–79.
52. Hermite, Sur la résolution de l'équation du cinquième degré, Comptes Rendus, 46, 1858, 508 ff. Also in his Œuvres, vol. II, Paris, 1908, 5–12.
53. Hermite, Sur quelques théorèmes d'algèbre et la résolution de l'équation du quatrième degré, Comptes Rendus, 46, 1858, 961 ff.; Œuvres, vol. II, 30–37.
54. Hermite, Sur l'invariant du 18e ordre des formes du cinquième degré, Crelle's Journal, 59, 1861, 304–305.
55. Hermite, Sur l'équation du cinquième degré, Comptes Rendus, 61 and 62, 1865 and 1866. The article is scattered throughout the two volumes. It is collected in his Œuvres, vol. II, 347–424.
56. Heymann, Über die Auflösung der Gleichungen vom fünften Grade, Zeits. für Math. und Phys., 39, 1894, 162–182.
57. Heymann, Die Transformation und Auflösung der Gleichung fünften Grades, Zeits. für Math. und Phys., 42, 1897, 81–98, and 113–121.
58. Hilbert, Über die Gleichung neunten Grades, Math. Annalen, 97, 1926, 243–250.
59. Hirsch, Sammlung von Aufgaben aus der Theorie der algebraischen Gleichungen, Berlin, 1809.
60. *Hunrath, Algebraische Untersuchungen nach Tschirnhausen's Methode, I. Progr. Glückstadt, 1876.
61. Jerrard, Mathematical Researches, Part. II, Bristol and London, 1834.
62. Jerrard, On a certain Transformation connected with the finite Solution of Equations of the Fifth Degree, Phil. Mag., ser. 3, 7, 1835, 202–203.
63. Jerrard, Reflections on the Resolution of Algebraic Equations of the Fifth Degree, Phil. Mag., ser. 3, 26, 1845, 545–574.
64. *Joubert, Sur l'équation du sixième degré, Comptes Rendus, 64, 1867.
65. Kiepert, Über die Auflösung der Gleichungen fünften Grades. Annali di Mat., ser. 2, 9, 1878–9, 119–123.
66. Kiepert, Auflösung der Gleichungen fünften Grades, Crelle's Journal, 87, 1879, 114–133.
67. Klein, Weitere Untersuchungen über das Ikosaeder, Math. Annalen, 12, 1877, 503–560.
68. Klein, Auflösung der Gleichungen fünften Grades, Math. Annalen, 14, 1879, 111–172.
69. Klein, Lectures on the Icosahedron, English translation by Morrice, 2nd ed., London, 1913. The English edition, however, is often unreliable.
70. *Kronecker, Über Gleichungen fünften Grades, Monatsber. d. Berl. Akademie, 1861, 609 ff.
71. Kronecker, Über die Gleichungen fünften Grades, Crelle's Journal, 59, 1861, 306–310.
72. Lagrange, Réflexions sur la résolution algébrique des équations, Nouv. Mémoires de l'Académie de Berlin, 1770–1. Also in his Œuvres, vol. III, Paris, 1869, 205–421.
73. Lavagna, Sulla determinazione dei coefficienti della trasformata di Tschirnhaus, Annali di Mat., ser. 1, 1, 1858, 238–249.
74. Leibniz, Briefwechsel, edited by Gerhardt, Berlin, 1899. Especially pages 315, 333, 359, 403, 449, 451.

75. *Maschke, La risoluzione della equazione del sesto grado, Rendiconti della R. Accademia dei Lincei, ser. 4, 4, 1888, 181–182.
76. Matthiessen, Grundzüge der antiken und modernen Algebra der litteralen Gleichungen, Leipzig, 1878. Page references are found in the index.
77. Netto, Vorlesungen über Algebra, vol. I, Leipzig, 1896, lecture 10.
78. Niewenglowski, Cours d'Algèbre, 5th ed., vol. II, Paris, 1902. Especially pages 303-333.
79. Petersen, Theorie der Algebraischen Gleichungen, Kopenhagen, 1878, 82–84.
80. Pincherle, Lezioni di Algebra, Bologna, 1909, 229 ff.
81. Rahts, Zur Reduction der allgemeinen Gleichung fünften Grades auf die Jerrard'sche Form, Math. Annalen, 28, 1886, 34-60.
82. Salmon, Modern Higher Algebra, 4th ed., Dublin, 1885, 250 ff.
83. Schlömilch, Die Transformation und Auflösung der Gleichungen fünften Grades nach Jerrard und Hermite, Zeits. für Math. und Phys., 4, 1859, 77–90.
84. Serret, Cours d'Algèbre Supérieure, 5th ed., vol. I, Paris, 1885. Especially page 132 ff. and 424 ff.
85. Sievers, Entwicklung der Umformung der Gleichung $z^5 - a_0 z^4 + b_0 z^3 - c_0 z^2 + d_0 z - e_0 = 0$ in $y^5 + y + \lambda = 0$, mittelst Tschirnhausen'schen Substitutionen, Astron. Nachr., 70, 1868, 355–358.
86. Simonart, Sur une transformation homographique d'une équation, Mathesis, 36, 1922, 212–218.
87. Sylvester, On the so-called Tschirnhausen Transformation, Crelle's Journal, 100, 1887, 465–486. Also in his Math. Papers, vol. IV, Cambridge, 1912, 531–549.
88. Sylvester, Sur une découverte de M. James Hammond, Comptes Rendus, 104, 1887, 1228–1231; Papers, vol. IV, 550–552.
89. Sylvester and Hammond, On Hamilton's Numbers, Phil. Trans., 178, 1887, 285–312, and 179, 1888, 65–71; Papers, vol. IV, 553–584.
90. *Tegman, Regula Cardani et methodus Bringiani radices inveniendi cubicas inter se collatae, Lund, 1798.
91. *Tegman, De aequatione biquadratica, Lund, 1799.
92. Todhunter, Theory of Equations, London, 1888, chapter XXVII.
93. Tortolini, Sulla risoluzione algebrica dell'equazioni di terzo e quarto grado, Annali di Mat., ser. 1, 1, 1858, 310–322.
94. *Tschirnhaus, Methodus auferendi omnes terminos intermedios ex data aequatione, Acta Eruditorum, 2, 1683, 204–207. The work of Tschirnhaus is essentially reproduced in references 23, 74 and 98 of this bibliography.
95. Tychsen, Om Transformationen af den almindelige Ligning af 5te Grad, Math. Tidsskrift, 5, 1863, 160–164.
96. Weber, Lehrbuch der Algebra, 2nd ed., vol. I, Braunschweig, 1898, chapters IV and VI.
97. *Weissenborn, Lebensbeschreibung des Ehrenfr. Walther v. Tschirnhaus auf Kiesslingswalde und Würdigung seiner Verdienste, Eisenach, 1866.
98 Weissenborn, Beiträge zu der Lehre von der Transformation der Gleichungen, Eisenach, 1870.
99. Wiman, Über gewisse imprimitive Gleichungen, Arkiv för Matematik, Astronomie, och Fysik, 3, 1906–07, No. 28.
100. Zeuthen, Geschichte der Mathematik im XVI. und XVII. Jahrhundert, Leipzig, 1903. Page references are found in the index.

ON THE FACTORING OF FREDHOLM MINORS.*

By H. T. Davis.

1. Introduction. The striking similarity between a certain class of theorems in algebra and the theory of the integral equation

$$u(x) = f(x) + \lambda \int_a^b K(x, t)\, u(t)\, dt, \tag{1}$$

has often been pointed out. Fredholm's first work† and the subsequent developments of D. Hilbert‡ showed that the Fredholm theory may be regarded as the limiting case of a system of m linear equations in n variables where m and n increase without limit. J. Mercer§ in discussing functions of positive and negative type in their relation to integral equations made frequent appeal to the algebraic theory of quadratic forms as a guide to his intuitions.

The object of the present study is to point out a significant connection between the algebraic theory of bilinear forms and the theory of the transcendental problem presented by the Fredholm equation as it relates to the factoring of Fredholm minors. Incidentally new identities, equations (14) and (19), associated with these minors are developed.

By an ingenious but somewhat indirect method L. Tocchi‖ has proved the following theorem:

The Fredholm minor associated with the integral equation (1),

$$D\begin{pmatrix} x \\ y \end{pmatrix} = \sum_{m=0}^{\infty} (-1)^m A_m \begin{pmatrix} x \\ y \end{pmatrix} \lambda^m, \tag{2}$$

where we use the abbreviations,

$$K\begin{pmatrix} x_1 x_2 \cdots x_m \\ y_1 y_2 \cdots y_m \end{pmatrix} = \begin{vmatrix} K(x_1, y_1), & K(x_1, y_2), & \cdots, & K(x_1, y_m) \\ K(x_2, y_1), & K(x_2, y_2), & \cdots, & K(x_2, y_m) \\ \cdots & \cdots & \cdots & \cdots \\ K(x_m, y_1), & K(x_m, y_2), & \cdots, & K(x_m, y_m) \end{vmatrix},$$

* Received November 2, 1928. Presented to the American Mathematical Society, April 16, 1927.

† Sur une classe d'équations fonctionnelles. Acta Math., vol. 27 (1903), pp. 365–390.

‡ Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Leipzig, (1912), parts IV and V.

§ Functions of Positive and Negative Type and their Connection with the Theory of Integral Equations. Trans. Lond. Phil. Soc., (A), vol. 209 (1909), pp. 415–446.

‖ Sopra una classe d'equazioni integrali. Per. di Mat., (3), vol. 12 (1915), pp. 253–261; also Due Teoremi sulle serie di Fredholm. Giorn. di Mat., vol. 54 (1916), pp. 141–150.

$$(3) \qquad A_m\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{m!}\int_a^b \cdots \int_a^b K\begin{pmatrix} x\, x_1\, x_2 \cdots x_m \\ y\, x_1\, x_2 \cdots x_m \end{pmatrix} dx_1\, dx_2 \cdots dx_m,$$

is factorable into the product of a function of x by a function of y provided λ is a root of the equation,

$$(4) \qquad D(\lambda) = \sum_{m=0}^{\infty} (-1)^m A_m \lambda^m = 0,$$

where we mean,

$$(5) \qquad A_m = \frac{1}{m!}\int_a^b \cdots \int_a^b K\begin{pmatrix} x_1\, x_2 \cdots x_m \\ x_1\, x_2 \cdots x_m \end{pmatrix} dx_1\, dx_2 \cdots dx_m.$$

The generalization of this theorem to the case of Fredholm minors of any order was also made by Tocchi.

The present paper obtains Tocchi's results by applying to the problem the following simple criterion for the factoring of a function of two variables.

LEMMA. *If $D(x, y)$ is a function of two variables and possesses first and second derivatives with respect to both, then a necessary and sufficient condition that it be factorable into a function of x and a function of y, $D(x, y) = u(x)\, v(y)$, is that it satisfy the following equation:*

$$(6) \qquad \Delta(D) \equiv D(x, y)\frac{\partial^2 D}{\partial x\, \partial y} - \frac{\partial D}{\partial x}\frac{\partial D}{\partial y} = 0.$$

That this condition is necessary is at once obvious. The sufficiency of the condition follows from the fact that (6) can be put into the form:

$$D^2(x, y)\frac{\partial}{\partial y}\left[\frac{\partial D}{\partial x}\Big/ D\right] = 0,$$

provided $D(x, y)$ is not identically zero. The trivial nature of the theorem when this condition is not satisfied, however, shows that no essential restriction has been imposed by this assumption.

Integrating we get $\dfrac{\partial D}{\partial x}\Big/ D = \varphi(x)$, where $\varphi(x)$ is an arbitrary function of x. Integrating again we obtain finally,

$$\log D = \int^x \varphi(x)\, dx + \psi(y),$$

where $\psi(y)$ is arbitrary, and hence,

$$D = e^{\int^x \varphi(x)\, dx}\, e^{\psi(y)}.$$

2. The factoring of bilinear forms. The lemma of the preceding section can be used to derive criteria for the factoring of bilinear forms in the form in which they will apply most usefully to the theory of the present paper.

Considering the simplest case, let us write the bilinear form as the determinant,

$$ {}_nD_1 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ y_1 & y_2 & \cdots & y_n & 0 \end{vmatrix}, $$

in which we shall assume that the x_i are functions of a single parameter x and the y_i functions of a second independent parameter y. The matrix of the form is evidently,

$$ A = -\begin{Vmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{Vmatrix}, $$

where the A_{ij} are the cofactors of a_{ij} in the determinant

$$ a = |a_{ij}|, $$

which we shall assume is of rank not lower than $n-1$.

Applying the criterion of the lemma to D which we can write in the form

$$ D = -\sum_{ij=1}^{n} A_{ij}\, x_i\, y_j, $$

we obtain the following expansion:

$$ \begin{aligned} \Delta(D) &= -\left(\sum_{ij=1}^{n}\sum_{kl=1}^{n} A_{ij}\, x_i\, y_j\, A_{kl}\, x'_k\, y'_l - \sum_{ij=1}^{n}\sum_{kl=1}^{n} A_{ij}\, x'_i\, y_j\, A_{kl}\, x_k\, y'_l\right), \\ &= -\sum_{ij=1}^{n}\sum_{kl=1}^{n} (A_{ij}\, A_{kl} - A_{kj}\, A_{il})\, x_i\, y_j\, x'_k\, y'_l, \\ &= -\sum_{ij=1}^{n}\sum_{kl=1}^{n} \begin{vmatrix} A_{ij} & A_{il} \\ A_{kj} & A_{kl} \end{vmatrix} x_i\, y_j\, x'_k\, y'_l. \end{aligned} $$

But since any second order determinant of the adjoint of $a = |a_{ij}|$ contains a as a factor, we have,

$$\Delta(D) = -\sum_{ij=1}^{n}\sum_{kl=1}^{n}\begin{vmatrix} A_{ij} & A_{il} \\ A_{kj} & A_{kl} \end{vmatrix} x_i\, y_j\, x'_k\, y'_l,$$

(7)

$$= -\,a\sum_{ij=1}^{n}\sum_{kl=1}^{n} A_{\binom{ik}{jl}}\, x_i\, y_j\, x'_k\, y'_l,$$

where $A_{\binom{ik}{jl}}$ is the algebraic complement of the determinant $\begin{vmatrix} a_{ij} & a_{il} \\ a_{kj} & a_{kl} \end{vmatrix}$.

But we also know that $\sum_{ij=1}^{n}\sum_{kl=1}^{n} A_{\binom{ik}{jl}}\, x_i\, y_j\, x'_k\, y'_l$ is equivalent to the expansion of the determinant:

$$b = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 & x'_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 & x'_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n & x'_n \\ y_1 & y_2 & \cdots & y_n & 0 & 0 \\ y'_1 & y'_2 & \cdots & y'_n & 0 & 0 \end{vmatrix}.$$

Hence equation (7) becomes

$$\Delta(D) = -\,a \cdot b, \tag{8}$$

and thus we find that $\Delta(D) \equiv 0$ if either a or b is zero. In the latter case, however, we should have $A_{\binom{ik}{jl}} = 0$, which is contrary to our assumption that a is of rank not lower than $n-1$. Thus we have proved that D factors into the product of a linear form of the x_i by a linear form of the y_i if and only if $a = 0$.*

With more generality let us consider the form represented by the expansion of the determinant:

$$_nD_p = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_{11} & x_{12} & \cdots & x_{1p} \\ a_{21} & a_{22} & \cdots & a_{2n} & x_{21} & x_{22} & \cdots & x_{2p} \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_{n1} & x_{n2} & \cdots & x_{np} \\ y_{11} & y_{12} & \cdots & y_{1n} & 0 & 0 & \cdots & 0 \\ y_{21} & y_{22} & \cdots & y_{2n} & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ y_{p1} & y_{p2} & \cdots & y_{pn} & 0 & 0 & \cdots & 0 \end{vmatrix}, \quad p \leqq n, \tag{9}$$

where the determinant $a = |a_{ij}|$ is of rank not lower than $n-p$.

* This condition, it should be observed, is well known and can be derived without use of the lemma, which has been introduced because of its subsequent usefulness. Thus we might argue that, since D is factorable, $|A_{ij}|$ is of rank $\leqq 1$ and $\begin{vmatrix} A_{ij} & A_{ik} \\ A_{lj} & A_{lk} \end{vmatrix} = a \cdot A_{\binom{ij}{kl}} = 0$. But $|a_{ij}|$ is of rank $\geqq n-1$. Hence some $A_{\binom{ij}{kl}} \neq 0$, and, therefore, $a = 0$.

If we regard this as a bilinear form in the variables $x_{1i}, x_{2i}, \cdots, x_{ni}$, which depend upon a parameter x, and of $y_{i1}, y_{i2}, \cdots, y_{in}$, which depend upon a parameter y, we can apply the previous result and thus obtain,

$$\Delta({}_nD_p) = -a_i \cdot b_i, \qquad (i = 1, 2, 3, \cdots, p),$$

where a_i is the determinant ${}_nD_p$ with the $(n+i)$th row and column deleted and b_i is the determinant ${}_nD_p$ bordered with a new column $x'_{1i}, x'_{2i}, \cdots, x'_{ni}$ and a new row $y'_{i1}, y'_{i2}, \cdots, y'_{in}$.

A necessary and sufficient condition for the vanishing of $\Delta({}_nD_p)$ and hence for the factoring of ${}_nD_p$ into the product of a linear form in $x_{1i}, x_{2i}, \cdots, x_{ni}$ by a linear form in $y_{1i}, y_{2i}, \cdots, y_{ni}$, is thus seen to be

$$a_i = 0,$$

where i may have any of the values from 1 to p.

3. **The bilinear kernel.** In order to exhibit the connection between the algebraic and the transcendental problem, we first consider the case where the kernel of the integral equation (1) is of the form,

$$K(x, y) = \sum_{i=1}^{n} X_i(x)\, Y_i(y). \tag{10}$$

It is important to note that a general kernel $K(x, y)$ can be represented by the series,

$$K(x, y) = \sum_{i=1}^{n} u_i(x)\, v_i(y)/\lambda_i, \tag{11}$$

provided it is uniformly convergent, where λ_i are the roots of $D(\lambda) = 0$, and $\{u_i(x)\}$, $\{v_i(y)\}$, the system of adjoint normalized biorthogonal functions associated with the kernel $K(x, y)$ and the characteristic numbers $\{\lambda_i\}$.* Thus we are justified *heuristically* in assuming that theorems that apply to the bilinear kernel (10) apply in the limit to (11) and hence to the general kernel.

If the kernel is of form (10), then $D(\lambda)$ and $D\begin{pmatrix} x \\ y \end{pmatrix}$ reduce to the following determinants:

$$D(\lambda) = |a_{ij}|, \qquad D\begin{pmatrix} x \\ y \end{pmatrix} = -{}_nD_1,$$

* See E. Schmidt: Zur Theorie der linearen und nichtlinearen Integralgleichungen. Math. Ann., vol. 63 (1907), pp. 433–476; in particular, p. 466.

E. Goursat: Sur un cas élémentaire de l'équation de Fredholm. Bull. de la soc. math. de France, vol. 35 (1907), pp. 163–173; H. Lebesgue: Sur la méthode de Goursat pour la résolution de l'équation de Fredholm. Ibid., vol. 36 (1908), pp. 3–19.

where the elements become, $x_i = X_i(x)$, $y_i = Y_i(y)$,

$$a_{ij} = \delta_{ij} - \lambda \int_a^b X_i(t)\, Y_j(t)\, dt, \qquad \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}. \tag{12}$$

It is obvious that the theory of the preceding section now applies and the first Fredholm minor factors into the product of a linear form of the $X_i(x)$ by a linear form of the $Y_i(y)$ provided $D(\lambda) = 0$.

For the bilinear kernel the Fredholm minor of pth order takes the form:

$$\begin{aligned} D\begin{pmatrix} x_1 x_2 \cdots x_p \\ y_1 y_2 \cdots y_p \end{pmatrix} &= \sum_{m=0}^{n} (-1)^m A_m \begin{pmatrix} x_1 x_2 \cdots x_p \\ y_1 y_2 \cdots y_p \end{pmatrix} \lambda^m, \\ &= (-1)^p {}_nD_p, \end{aligned} \tag{13}$$

where the elements of (9) become $x_{ij} = X_i(x_j)$, $y_{ji} = Y_i(y_j)$,

$$a_{ij} = \delta_{ij} - \lambda \int_a^b X_i(t)\, Y_j(t)\, dt, \qquad \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

This fact, as far as the author is aware, has never been explicitly pointed out except for the case where $p = 1$. The proof involves a long and tedious algebraic development following the lines indicated in the paper of Goursat to which reference has already been made. The result, however, can be derived by identifying a formula given by Tocchi in the first paper cited above with the formal development of the bordered determinant (13). Since the algebraic details are involved and add nothing to the discussion of the present problem they will be omitted.*

Applying the results of the preceding section to (13) and differentiating with respect to x_i and y_i, we get:

$$\Delta\left\{D\begin{pmatrix} x_1 x_2 \cdots x_p \\ y_1 y_2 \cdots y_p \end{pmatrix}\right\} = D\begin{pmatrix} x_1 x_2 \cdots x_{i-1} x_{i+1} \cdots x_p \\ y_1 y_2 \cdots y_{i-1} y_{i+1} \cdots y_p \end{pmatrix} D_i', \qquad (i = 1, 2, \cdots, p)$$

where D_i' is understood to be the determinant (13) bordered by the derived quantities $X_1'(x_i)$, $X_2'(x_i)$, $\cdots$, $X_n'(x_i)$ and $Y_1'(y_i)$, $Y_2'(y_i)$, $\cdots$, $Y_n'(y_i)$.

It thus appears that $D\begin{pmatrix} x_1 x_2 \cdots x_p \\ y_1 y_2 \cdots y_p \end{pmatrix}$ factors into a product of a linear form of the $X_k(x_i)$, $k = 1, 2, \cdots, n$, by a linear form of the $Y_k(y_i)$ provided we have, $D\begin{pmatrix} x_1 x_2 \cdots x_{i-1} x_{i+1} \cdots x_p \\ y_1 y_2 \cdots y_{i-1} y_{i+1} \cdots y_p \end{pmatrix} = 0$.

4. **The transcendental case.** The simplicity of the approach to the general case by the method developed above is found in the easy generali-

* Compare formula (12) in the first paper of Tocchi with the development given by G. Kowalewski, Determinantentheorie, (1909), pp. 90–92.

zation from the algebraic to the transcendental problem. What we can say about the determinant ${}_nD_p$ applies to the limit as $n \to \infty$ provided $\sum_{i=1}^{\infty} X_i(s)\, Y_i(t)$ converges uniformly for all values of s and t in the rectangle $a \leqq s \leqq b,\ a \leqq t \leqq b$.

Thus equation (7) is seen to reduce to the following:

$$\Delta(D) = \left\{\sum_{m=0}^{\infty} (-1)^m \lambda^m A_m\right\} \times \left\{\sum_{m=0}^{\infty} (-1)^m \frac{\lambda^m}{m!} \int_a^b \cdots \int_a^b S_m \begin{pmatrix} x \\ y \end{pmatrix} d t_1\, d t_2 \cdots d t_m\right\}, \tag{14}$$

where we use the abbreviation:

$$S_m \begin{pmatrix} x_1\ x_2 \cdots x_p \\ y_1\ y_2 \cdots y_p \end{pmatrix} = \begin{vmatrix} K''_{x_1 y_1}(x_1, y_1) & K'_x(x_1, y_1) \cdots & K'_x(x_1, y_p) & K'_x(x_1, t_1) \cdots & K'_x(x_1, t_m) \\ K'_{y_1}(x_1, y_1) & K(x_1, y_1) \cdots & K(x_1, y_p) & K(x_1, t_1) \cdots & K(x_1, t_m) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ K'_{y_1}(x_p, y_1) & K(x_p, y_1) \cdots & K(x_p, y_p) & K(x_p, t_1) \cdots & K(x_p, t_m) \\ K'_{y_1}(t_1, y_1) & K(t_1, y_1) \cdots & K(t_1, y_p) & K(t_1, t_1) \cdots & K(t_1, t_m) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ K'_{y_1}(t_m, y_1) & K(t_m, y_1) \cdots & K(t_m, y_p) & K(t_m, t_1) \cdots & K(t_m, t_m) \end{vmatrix}.$$

If we denote by M a number greater than or equal to the upper bounds of $|K''_{xy}(x, y)|$, $|K'_x(x, y)|$, $|K'_y(x, y)|$, and $|K(x, y)|$ over the rectangle of definition, then it follows from Hadamard's theorem* that $\left|S_m \begin{pmatrix} x \\ y \end{pmatrix}\right| \leqq (m+2)^{(m+2)/2} M^{m+2}$. Since, moreover, we have

$$\lim_{m=\infty} \sqrt[m]{(m+2)^{(m+2)/2} M^{m+2}/m!} = 0,$$

the second factor of (14) is seen to converge for all values of λ.

Because of the method of derivation of (14) as the limiting case of an algebraic identity, it is important to establish this equation directly. We can accomplish this by considering the following two expansions for $D\begin{pmatrix} x \\ y \end{pmatrix}$:†

$$D \begin{pmatrix} x \\ y \end{pmatrix} = D(\lambda)\, \{a_0 + a_1 \lambda + a_2 \lambda^2 + \cdots\}, \tag{15}$$

where $a_0 = K\begin{pmatrix} x \\ y \end{pmatrix}$ and $a_m = \int_a^b \int_a^b K(x, t)\, K(t_1, t_2) \cdots K(t_m, y)\, d t_1 \cdots d t_m$; and

* Bull. des Sci. Math., vol. 17 (1893), pp. 240–246.

† See E. Goursat: Cours d'analyse, vol. 3 (1915), p. 346 and p. 371.

$$(16) \qquad D\begin{pmatrix} x \\ y \end{pmatrix} = b_0 + b_1\lambda + \frac{b_2}{2!}\lambda^2 + \frac{b_3}{3!}\lambda^3 + \cdots,$$

where

$$(17) \quad b_0 = K\begin{pmatrix} x \\ y \end{pmatrix} \text{ and } b_m = (-1)^m \int_a^b \cdots \int_a^b K\begin{pmatrix} x\, t_1\, t_2 \cdots t_m \\ y\, t_1\, t_2 \cdots t_m \end{pmatrix} dt_1\, dt_2 \cdots dt_m.$$

Let us form the expression by direct multiplication of these two expansions, using (15) for the terms of $\partial D/\partial x$ and $\partial^2 D/\partial x\, \partial y$ and (16) for the terms of D and $\partial D/\partial y$.

The coefficient of $\lambda^r/r!$ is seen to be

$$(18) \qquad \sum_{i=0}^{r} (-1)^i \left\{ b_{r-i} \frac{\partial^2 a_i}{\partial x\, \partial y} - \frac{\partial b_{r-i}}{\partial y} \frac{\partial a_i}{\partial x} \right\} \frac{r!}{(r-i)!}.$$

Since each b_{r-i} consists of $(r-i+1)!$ terms it will be seen that (18) contains $2\sum_{i=0}^{r} (r-i+1)!\, r!/(r-i)! = r!\,(r+1)\,(r+2) = (r+2)!$ terms. This is precisely the same number that we have in $\int_a^b \cdots \int_a^b S_r\begin{pmatrix} x \\ y \end{pmatrix} dt_1 \cdots dt_r$. The identification of terms is made as follows. The kth term of (18) may be explicitly written:

$$\begin{aligned} I_k = (-1)^k \frac{r!}{(r-k)!} \Bigg\{ & \int_a^b \cdots \int_a^b K\begin{pmatrix} x\, t_1\, t_2 \cdots t_{r-k} \\ y\, t_1\, t_2 \cdots t_{r-k} \end{pmatrix} dt_1 \cdots dt_{r-k} \\ & \times \int_a^b \cdots \int_a^b \frac{\partial K(x, t_1)}{\partial x} K(t_1, t_2) \cdots \frac{\partial K(t_k, y)}{\partial y}\, dt_1 \cdots dt_k \\ & - \int_a^b \cdots \int_a^b \frac{\partial K}{\partial y}\begin{pmatrix} x\, t_1\, t_2 \cdots t_{r-k} \\ y\, t\ \ t \cdots t_{r-k} \end{pmatrix} dt_1 \cdots dt_{r-k} \\ & \times \int_a^b \cdots \int_a^b \frac{\partial K(x, t)}{\partial x} K(t_1, t_2) \cdots K(t_k, y)\, dt_1 \cdots dt_k \Bigg\}. \end{aligned}$$

In the second integral of each product change the subscripts of each integration variable, letting these subscripts range from $r-k+1$ to r. We can then write I_k in the form

$$\begin{aligned} I_k = (-1)^k \frac{r!}{(r-k)!} \int_a^b \cdots \int_a^b \Bigg\{ & K\begin{pmatrix} x\, t_{s_1}\, t_{s_2} \cdots t_{s_{r-k}} \\ y\, t_{s_1}\, t_{s_2} \cdots t_{s_{r-k}} \end{pmatrix} K_x'(x, t_{r_1})\, K(t_{r_1}, t_{r_2}) \cdots \\ & \cdots K_y'(t_{r_k}, y) - K_y'\begin{pmatrix} x\, t_{s_1}\, t_{s_2} \cdots t_{s_{r-k}} \\ y\, t_{s_1}\, t_{s_2} \cdots t_{s_{r-k}} \end{pmatrix} K_x'(x, t_{r_1})\, K(t_{r_1}, t_{r_2}) \cdots \\ & \cdots K(t_{r_k}, y) \Bigg\}\, dt_1\, dt_2 \cdots dt_r, \end{aligned}$$

where $s_1, s_2, \cdots, s_{r-k}$ is any permutation of the numbers $1, 2, \cdots, r-k$, and $r_1, r_2, \cdots, r_k$ any permutation of the numbers $r-k+1, \cdots, r$. If, further, the r_j are chosen in any order from the numbers $1, 2, \cdots, r$ and such changes made in corresponding values of the s_i so that all of the integers from 1 to r are represented in the sequences $s_1, s_2, \cdots, s_{r-k}$; $r_1, r_2, \cdots, r_k$, then we shall have $r!/(r-k)!$ separate representations of the integrand, each one of which is seen to be a term in the expansion of $S_r\begin{pmatrix} x \\ y \end{pmatrix}$. Thus in the expansion of $S_r\begin{pmatrix} x \\ y \end{pmatrix}$ the determinant $K\begin{pmatrix} x\, t_{s_1} \cdots t_{s_{r-k}} \\ y\, t_{s_1} \cdots t_{s_{r-k}} \end{pmatrix}$ is multiplied by its complement. The terms in the latter corresponding to $K'_x(x, t_{r_1})\, K(t_{r_1}, t_{r_2}) \cdots K'_y(t_{r_k}, y)$ are seen from the cyclical nature of the subscripts of t, to be those having the sign $(-1)^k$. Since in the complete expansion of (18) none of these terms repeat themselves, a unique correspondence between the terms of (18) and the terms of $S_r\begin{pmatrix} x \\ y \end{pmatrix}$ is thus established.

From (18) we derive the theorem that $D\begin{pmatrix} x \\ y \end{pmatrix}$ factors into a function of x by a function of y provided,

$$D(\lambda) = 0.$$

Similarly for the general case where the derivatives in the symbol Δ apply to x_i and y_i, we have,

$$\Delta\left\{D\begin{pmatrix} x_1\, x_2 \cdots x_p \\ y_1\, y_2 \cdots y_p \end{pmatrix}\right\} = \left\{\sum_{m=0}^{\infty} (-1)^m \lambda^m A_m \begin{pmatrix} x_1\, x_2 \cdots x_{i-1}\, x_{i+1} \cdots x_p \\ y_1\, y_2 \cdots y_{i-1}\, y_{i+1} \cdots y_p \end{pmatrix}\right\} \times \left\{\sum_{m=0}^{\infty} (-1)^m \frac{\lambda^m}{m!} \int_a^b \cdots \int_a^b S_m \begin{pmatrix} x_1\, x_2 \cdots x_p \\ y_1\, y_2 \cdots y_p \end{pmatrix} dt_1\, dt_2 \cdots dt_m\right\}, \tag{19}$$

where we abbreviate,

$$A_m \begin{pmatrix} x_1\, x_2 \cdots x_p \\ y_1\, y_2 \cdots y_p \end{pmatrix} = \frac{1}{m!} \int_a^b \cdots \int_a^b K \begin{pmatrix} x_1\, x_2 \cdots x_p;\ t_1\, t_2 \cdots t_m \\ y_1\, y_2 \cdots y_p;\ t_1\, t_2 \cdots t_m \end{pmatrix} dt_1\, dt_2 \cdots dt_m.$$

Hence we deduce that $D\begin{pmatrix} x_1\, x_2 \cdots x_p \\ y_1\, y_2 \cdots y_p \end{pmatrix}$ will factor into a function of x_i by a function of y_i provided,

$$D\begin{pmatrix} x_1\, x_2 \cdots x_{i-1}\, x_{i+1} \cdots x_p \\ y_1\, y_2 \cdots y_{i-1}\, y_{i+1} \cdots y_p \end{pmatrix} \equiv \sum_{m=0}^{\infty} (-1)^m \lambda^m A_m \begin{pmatrix} x_1\, x_2 \cdots x_{i-1}\, x_{i+1} \cdots x_p \\ y_1\, y_2 \cdots y_{i-1}\, y_{i+1} \cdots y_p \end{pmatrix} = 0.$$

WATERMAN INSTITUTE, INDIANA UNIVERSITY.

A SET OF NECESSARY AND SUFFICIENT CONDITIONS FOR THE CESÀRO SUMMABILITY OF DOUBLE SERIES.*

By Gaylord M. Merriman.†

The main theorem of the present paper, though dealing entirely with general double series, is one whose importance is best illustrated in terms of double Fourier's series, to which it will be applied in a future paper. On account of the wide applicability of both single and double Fourier's series to boundary value problems, a close study of their summability, particularly of that kind defined by Cesàro, has been justified.‡ In view of the importance of this "Cesàro Summability Problem" and of its companion problem of convergence, it is a significant fact that the latter has never been completely solved, and that, until just recently, the theory relating to the former, as far as single series are concerned, has contained the same hiatus.§ It remained for Hardy and Littlewood|| to solve this latter problem for single Fourier's series, and that only in a modified form; thus they found conditions both necessary and sufficient for the Cesàro summability of a single Fourier's series, not with respect to a *given mean*, but with respect to *some mean or other*. This modification is justified by the simplicity of the final result.

Here we make a beginning toward solving the analogous problem for the double Fourier's series (i. e. toward finding necessary and sufficient conditions for the Cesàro summability of such series, with respect to *some pair of means or other*) by deducing a general result in the theory of double series. This theorem itself is interesting, and its importance may be realized from the number of applications which should be possible.¶

* Received October 21, 1927; presented to the American Mathematical Society, Chicago, April 1926.

† National Research Fellow.

‡ The preference to be given the process of summability over that of convergence has been pointed out often. Cf., for example, a paper by Professor C. N. Moore: "Applications of the theory of summability to developments in orthogonal functions", Bull. Amer. Math. Soc., vol. 25 (1919), pp. 258–276.

§ By a complete solution of the problem is meant the derivation of a condition that is both necessary and sufficient to insure the result desired.

|| Hardy and Littlewood: "Solution of the Cesàro summability problem for power series and Fourier series", Math. Zeit., vol. 19 (1923–24), pp. 67–96. We shall refer often to this paper by the initials H.-L.

¶ For some of the applications of the theory relating to single series cf. H.-L., II.6 and III; also, cf. another paper by the same authors: "The allied series of a Fourier series", Proc. Lond. Math. Soc., vol. 24 (1924–1925), pp. 211–246.

The deductions are, for the most part, generalizations to two dimensions of those concerning single series obtained by Hardy and Littlewood in the paper first cited above.

1. **Definitions and notation.** We have for consideration a double series

$$A = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{i,j}, \tag{1}$$

where A shall stand for either the series or its sum; by A_C and A_R we shall mean the simple series represented by any column or any row of (1), or the sum of such a series. We write also, for convenience,

$$\begin{aligned} A_{m,n}^{-1,-1} &= a_{m,n}, \quad A_{m,n}^{-1,0} = \sum_{j=0}^{n} a_{m,j}, \quad A_{m,n}^{0,-1} = \sum_{i=0}^{m} a_{i,n}, \\ A_{m,n}^{0,0} &= \sum_{0}^{m} \sum_{0}^{n} a_{i,j} = A_{m,n}, \\ A_{m,n}^{0,1} &= \sum_{j=0}^{n} A_{m,j}^{0,0}, \quad A_{m,n}^{1,0} = \sum_{i=0}^{m} A_{i,n}^{0,0}, \\ A_{m,n}^{1,1} &= \sum_{0}^{m} \sum_{0}^{n} A_{i,j}^{0,0}, \quad \text{etc.} \end{aligned} \tag{2}$$

The series (1) is said to be summable (C, r, s) to sum A if

$$A_{m,n}^{r,s} \sim C_{m,n}^{r,s} A,$$

where

$$C_{m,n}^{r,s} = \frac{(m+r)(m+r-1)\cdots(m+1)(n+s)(n+s-1)\cdots(n+1)}{r!\, s!}.$$

In these circumstances we say that $A_{m,n} \to A$ (C, r, s). Analogously, we shall use the notations

$$\begin{aligned} A_{m,n} &\to A_R \ (C, s) \qquad \text{(every } m\text{)}, \\ A_{m,n} &\to A_C \ (C, r) \qquad \text{(every } n\text{)}, \end{aligned}$$

to indicate Cesàro summability of rows or columns of (1), considered as simple series. When $r = s$, we shall say that

$$A_{m,n} \to A \ (C, r),$$

provided that such a notation implies no ambiguity. When (1) is summable by some pair of unspecified means, we shall say that it is summable (C); and we shall use the same terminology to indicate summability of A_R or A_C with respect to some mean or other.

Summability $(C, 0, 0)$ is the same as convergence. We say that A is summable $(C, -1, -1)$ if (i) it is convergent, (ii) $A_{m,n}^{-1,0}$ is $o(1/m)$, and $A_{m,n}^{0,-1}$

is $o(1/n)$, and (iii) $a_{m,n}$ is $o(1/mn)$. Again, A is summable $(C, r, -1)$ if (i) it is summable $(C, r, 0)$, and (ii) $A_{m,n}^{0,-1}$ is $o(1/n)$. A similar definition gives the meaning of summability $(C, -1, s)$. We shall use the definitions of summability $(C, -1)$ of A_R and A_C as given in H.-L.

Finally, we suppose that (1) satisfies the following conditions:

$$\begin{aligned} &\lim_{m,n\to\infty} a_{m,n} = 0, \\ &\lim_{m\to\infty} a_{m,n} = 0 \qquad \text{(for every } n\text{)}, \\ &\lim_{n\to\infty} a_{m,n} = 0 \qquad \text{(for every } m\text{)}. \end{aligned} \tag{3}$$

2. **Statement of the main theorem.** The main theorem of the paper, relating to the Cesàro summability of (1), is proved for the case $r = s$, since, in accordance with the problem as outlined in the introduction, we are interested here only in the *summability* of the series, and not in the *least pair* of means by which it may be summable. The theorem is as follows:

THEOREM I: *A set of necessary and sufficient conditions that* (1), *satisfying* (3), *and that the simple series represented by the rows and columns of* (1), *should all be summable* (C, r), *is that there should exist systems of numbers* $a_{m,n}^{\varrho,\varrho}$ $(m, n = 1, 2, \cdots)$, $a_{m,n}^{\varrho,\varrho-1}$ *(every* n; $m = 1, 2, \cdots)$, $a_{m,n}^{\varrho-1,\varrho}$ *(every* m; $n = 1, 2, \cdots)$, $\varrho = 1, 2, \cdots, r+1$, *connected with the terms of* (1) *through*

$$\begin{cases} a_{m,n}^{0,0} = a_{m,n}, \\ a_{m,n}^{\varrho-1,\varrho-1} = (m+1)(n+1)\{a_{m,n}^{\varrho,\varrho} - a_{m+1,n}^{\varrho,\varrho} - a_{m,n+1}^{\varrho,\varrho} + a_{m+1,n+1}^{\varrho,\varrho}\}, \end{cases} \tag{4}$$

$$\left.\begin{cases} a_{m,n}^{0,0} = a_{m,n} \\ a_{m,n}^{\varrho-1,\varrho-1} = (m+1)\{a_{m,n}^{\varrho,\varrho-1} - a_{m+1,n}^{\varrho,\varrho-1}\} \end{cases}\right\} \quad \textit{(every } n\textit{)}, \tag{5}$$

$$\left.\begin{cases} a_{m,n}^{0,0} = a_{m,n} \\ a_{m,n}^{\varrho-1,\varrho-1} = (n+1)\{a_{m,n}^{\varrho-1,\varrho} - a_{m,n+1}^{\varrho-1,\varrho}\} \end{cases}\right\} \quad \textit{(every } m\textit{)}, \tag{6}$$

such that

$$A^{r+1,r+1} \equiv \sum_0^\infty \sum_0^\infty a_{m,n}^{r+1,r+1}, \tag{7}$$

$$A^{r+1,r} \equiv \sum_{m=0}^\infty a_{m,n}^{r+1,r} \qquad \textit{(every } n\textit{)}, \tag{8}$$

$$A^{r,r+1} \equiv \sum_{n=0}^\infty a_{m,n}^{r,r+1} \qquad \textit{(every } m\textit{)}, \tag{9}$$

are summable $(C, -1)$.

In these circumstances, we have

$$a_{m,n}^{\varrho,\varrho} = \sum_m^\infty \sum_n^\infty \frac{a_{\mu,\nu}^{\varrho-1,\varrho-1}}{(\mu+1)(\nu+1)}, \ (C, r-\varrho), \tag{10}$$

$$(11) \qquad a_{m,n}^{\varrho,\varrho-1} = \sum_{m}^{\infty} \frac{a_{\mu,n}^{\varrho-1,\varrho-1}}{\mu+1}, \qquad (C, r-\varrho), \text{ for every } n,$$

$$(12) \qquad a_{m,n}^{\varrho-1,\varrho} = \sum_{n}^{\infty} \frac{a_{m,\nu}^{\varrho-1,\varrho-1}}{\nu+1}, \qquad (C, r-\varrho), \text{ for every } m,$$

for $\varrho = 1, 2, \cdots, r+1$. *The series* $A^{\varrho,\varrho}$ *is summable* $(C, r-\varrho)$ *to the same value as* A, *while* $A^{\varrho-1,\varrho}$ *and* $A^{\varrho,\varrho-1}$ *are summable* $(C, r-\varrho)$ *to the same values as* A_R *and* A_C, *respectively, for* $\varrho = 1, 2, \cdots, r+1$.

To prove this theorem we shall need a series of lemmas.

3. The auxiliary series, B, P, and Q. We first introduce three auxiliary series, B, P, and Q, connected with A through the following definitions:

$$(13) \qquad p_{m,n} = (m+1)\,(b_{m,n} - b_{m+1,n}),$$

$$(13') \qquad a_{m,n} = (n+1)\,(p_{m,n} - p_{m,n+1}),$$

$$(14) \qquad q_{m,n} = (n+1)\,(b_{m,n} - b_{m,n+1}),$$

$$(14') \qquad a_{m,n} = (m+1)\,(q_{m,n} - q_{m+1,n}),$$

$$(15) \qquad a_{m,n} = (m+1)\,(n+1)\,(b_{m,n} - b_{m+1,n} - b_{m,n+1} + b_{m+1,n+1}).$$

From the fact that (15) can be obtained through combination either of (13) and (13′), or of (14) and (14′), we see that these definitions are consistent.

LEMMA 1. *If* A, A_C, *and* A_R *are summable* (C, r), *then* (i) Q_C *and* P_R *are summable* $(C, r-1)$, (ii) P_C *and* Q_R *are summable* (C, r), *and* (iii) B_C *and* B_R *are summable* $(C, r-1)$.

(i) To prove (i), we have that, in view of (13′), m fixed, and (14′), n fixed, we can, through the medium of H.-L., Lemma 5, choose the p's and q's so that

$$(16) \qquad p_{m,n} = \sum_{\nu=n}^{\infty} \frac{a_{m,\nu}}{\nu+1}, \quad (C, r-1), \text{ for every } m,$$

$$(16') \qquad q_{m,n} = \sum_{\mu=m}^{\infty} \frac{a_{\mu,n}}{\mu+1}, \quad (C, r-1), \text{ for every } n.$$

On account of this choice we can further infer from the same source that P_R sums to A_R, and Q_C to A_C.

(ii) The proof of (ii) is more complicated. In the first place, we can infer that if A, A_C, and A_R are summable (C, r), then the series $\sum\sum(a_{\mu,\nu}/\nu+1)$ and $\sum\sum(a_{\mu,\nu}/\mu+1)$ are summable $(C, r, r-1)$ and $(C, r-1, r)$ respectively,

and their rows and columns are summable of corresponding orders.* Hence we further infer that either of these series has row-column or column-row summability, whether we start at the first column or row, or at a later one;† that is to say, the repeated sums

$$\sum_{m=0}^{\infty}\left(\sum_{n}\frac{a_{m,n}}{n+1}\right), \qquad \sum_{n=0}^{\infty}\left(\sum_{m}\frac{a_{m,n}}{m+1}\right), \text{ etc.,}$$

exist in the Cesàro sense. Now, starting with (16) and (16′), we find that

$$\sum_{m=0}^{\infty} p_{m,n} = \sum_{m=0}^{\infty}\left(\sum_{n}\frac{a_{m,n}}{n+1}\right)$$

exists in the Cesàro sense, as does $\sum_{n=0}^{\infty} q_{m,n}$, and (ii) is proved.

(iii) The proof of (iii) is similar to that of (i), if we use (13) and (14) and (ii).

Lemma 1 is therefore completely established.

LEMMA 2. *If A, A_C, and A_R are summable (C, r) to zero,‡ then*

$$P_{m,n+1}^{r,r-1} = C_m^r\, o(n^{r-1}), \quad \textit{uniformly,} \tag{17}$$

$$Q_{m+1,n}^{r-1,r} = C_n^r\, o(m^{r-1}), \quad \textit{uniformly.} \tag{17′}$$

In view of (13′) we have

$$\frac{A_{m,n}^{r,-1}}{C_m^r} = (n+1)\left[\frac{P_{m,n}^{r,-1}}{C_m^r} - \frac{P_{m,n+1}^{r,-1}}{C_m^r}\right]. \tag{18}$$

With (18) as a basic relation (comparable to (2.11) of H.-L.) it can be inferred from Lemmas 3 and 5 of H.-L. that

$$P_{m,n}^{r,r-1}/C_m^r = o(n^{r-1}),$$

whence (17) follows. (17′) is obtained in a similar manner. The uniformity of the limit processes in (17) and (17′) follows from the summability of A as a double series.

* This is a special case of Theorem IV of the author's previous paper: "Concerning the Cesàro summability of double series of a certain type" Annals of Math., vol. 28 (1927), pp. 515–533.

† By results analogous to certain ones of Pringsheim regarding convergence; cf. T. J. I'A. Bromwich: "Infinite Series", p. 75; also, cf. A. Pringsheim: "Vorlesungen über Zahlen- und Funktionenlehre", I. 2, pp. 449 ff.

‡ That we can take these sums all zero without loss of generality follows thus: In our given series we can first make the sum of each row zero by altering the elements of the first column to $\bar{a}_{m,0}$; then the sum of each column can be made zero by changing the elements of the first row to $\bar{a}_{0,n}$ and $\bar{a}_{0,0}$ to $\bar{\bar{a}}_{0,0}$; although this last change might seem to have affected the summability of the first row to zero, we infer that such is not the case because the sum of the whole series and those of all the other rows are zero.

LEMMA 3. *If A, A_C, and A_R are summable (C, r) to zero, then the sum of each of the various series mentioned in Lemma* 1 *is zero.*

(i) The proof of the summability $(C, r-1)$ to zero of Q_C and P_R is contained in (i) of Lemma 1.

(ii) For the proof of the summability to zero of P_C and Q_R we argue as follows: Because of the summability to zero of A_C and A_R we can conclude from (13′) and (14′) that successive columns of P and successive rows of Q sum to the same quantity. Next we conclude from Lemma 2 that P and Q are summable as double series to zero. Therefore, as P and Q sum as double series and also have row and column summability, we infer that repeated sums must give the same result as summation as a double series. Hence, because of the remark first made, the independent columns of P and the independent rows of Q must sum to zero.

(iii) The proof of the summability to zero of B_C and B_R follows the lines of the argument in (ii), if we use the results of Lemma 1 and (i) of the present lemma.

LEMMA 4. *If equations* (13)–(15) *are satisfied, then*

$$(19)\quad P^{r,r}_{m,n} = (r+1)\,B^{r,r}_{m,n} - (m+1)\,B^{r-1,r}_{m+1,n},$$

$$(19')\quad A^{r,r}_{m,n} = (r+1)\,P^{r,r}_{m,n} - (n+1)\,P^{r,r-1}_{m,n+1},$$

$$(20)\quad Q^{r,r}_{m,n} = (r+1)\,B^{r,r}_{m,n} - (n+1)\,B^{r,r-1}_{m,n+1},$$

$$(20')\quad A^{r,r}_{m,n} = (r+1)\,Q^{r,r}_{m,n} - (m+1)\,Q^{r-1,r}_{m+1,n},$$

$$(21)\quad A^{r,r}_{m,n} = (r+1)^2 B^{r,r}_{m,n} + (m+1)(n+1)\,B^{r-1,r-1}_{m+1,n+1} - (r+1)\left[(m+1)\,B^{r-1,r}_{m+1,n} + (n+1)\,B^{r,r-1}_{m,n+1}\right].$$

We shall indicate only the proofs of (20) and (21); the other results of the lemma will follow in the same way as (20).

In accordance with (14) we have, for the mth row of Q,

$$(22)\qquad Q^{-1,r}_{m,n} = (r+1)\,B^{-1,r}_{m,n} - (n+1)\,B^{-1,r-1}_{m,n+1},$$

from (2.12) of H.-L. Hence we obtain (20) by summation; and the other formulas (except (21)) follow analogously.

To prove (21), we first note that we can also obtain by summation of (22),

$$(23)\qquad Q^{r-1,r}_{m+1,n} = (r+1)\,B^{r-1,r}_{m+1,n} - (n+1)\,B^{r-1,r-1}_{m+1,n+1}.$$

Substitution of (23) and (20) in (20′) leads to (21).

4. Fundamental sufficiency lemma. LEMMA 5: *If B, B_C, and B_R are summable $(C, r-1)$ to zero, then A, A_R, and A_C are summable (C, r) to zero.*

We first note that, in view of (13) and (14), we can conclude that P_R and Q_C are summable $(C, r-1)$, since the difference of two simple series summable (C) to zero is also summable (C) to zero, and multiplication by a constant does not affect this property. Hence in view of (13′) and (14′) and the results of H.-L., Lemma 2, we can conclude that the summability $(C, r-1)$ of P_R and Q_C implies the summability (C, r) of A_R and A_C respectively, the sums all being zero. Therefore, to complete the lemma, we have only to concern ourselves with the proof of the results regarding the double series A and B.

(i) If $r = 0$, (21) becomes

$$A_{m,n} = B_{m,n} - (n+1)\sum_{i=0}^{m} b_{i,n+1} - (m+1)\sum_{j=0}^{n} b_{m+1,j} + (m+1)(n+1)\, b_{m+1,n+1}, \tag{24}$$

and the result follows at once, since the last three terms tend to zero as m and n become infinite.

(ii) If $r > 0$, we make the following reduction: By the hypothesis,

$$B^{r-1,r-1}_{m+1,n+1} = C^{r-1,r-1}_{m+1,n+1}\, B + o(m^{r-1}\, n^{r-1}). \tag{25}$$

By summing (25) once with respect to both indices, and to each index separately, we obtain

$$\begin{aligned} B^{r,r}_{m,n} &= C^{r,r}_{m,n}\ B + o(m^r\, n^r), \\ B^{r-1,r}_{m+1,n} &= C^{r-1\,r}_{m+1,n} B + o(m^{r-1}\, n^r), \\ B^{r,r-1}_{m,n+1} &= C^{r,r-1}_{m,n+1} B + o(m^r\, n^{r-1}). \end{aligned} \tag{26}$$

Substitution of (25) and (26) in (21) gives

$$A^{r,r}_{m,n} = \{(r+1)^2\, C^{r,r}_{m,n} - (r+1)[(m+1)\, C^{r-1,r}_{m+1,n} + (n+1)\, C^{r,r-1}_{m,n+1}] + (m+1)(n+1)\, C^{r-1,r-1}_{m+1,n+1}\}\, B + o(m^r\, n^r),$$

or

$$A^{r,r}_{m,n} = C^{r,r}_{m,n}\, B + o(m^r\, n^r),$$

and the conclusion follows at once.

5. **Fundamental necessity lemmas.** LEMMA 6. *If A, A_C, and A_R are summable (C, r) to sum zero, then*

$$B^{r-1,r-1}_{m,n} = h\, C^{r,r}_{m,n} + o(m^{r-1}\, n^{r-1}), \tag{27}$$

where h is a constant. If $r = 0$, (27) is to be interpreted as meaning

$$b_{m+1,n+1} = h + o(1/mn). \tag{28}$$

(i) We prove the lemma first when $r = 0$. In this case (21) yields

$$\begin{aligned}(29)\quad &(m+2)(n+2)B^{0,0}_{m,n} - (m+2)(n+1)B^{0,0}_{m,n+1} - (m+1)(n+2)B^{0,0}_{m+1,n} \\ &\qquad + (m+1)(n+1)B^{0,0}_{m+1,n+1} \\ &= B^{0,0}_{m,n} - (m+1)B^{-1,0}_{m+1,n} - (n+1)B^{0,-1}_{m,n+1} + (m+1)(n+1)b_{m+1,n+1} \\ &= A_{m,n} = o(1).\end{aligned}$$

Hence

$$\begin{aligned}(30)\quad &\frac{B^{0,0}_{m,n}}{(m+1)(n+1)} - \frac{B^{0,0}_{m,n+1}}{(m+1)(n+2)} - \frac{B^{0,0}_{m+1,n}}{(m+2)(n+1)} + \frac{B^{0,0}_{m+1,n+1}}{(m+2)(n+2)} \\ &\qquad = \frac{A^{0,0}_{m,n}}{(m+2)(m+1)(n+2)(n+1)} = o\left(\frac{1}{m^2n^2}\right),\end{aligned}$$

whence we obtain

$$\begin{aligned}(31)\quad &\frac{B^{0,0}_{m,n}}{(m+1)(n+1)} - \frac{B^{0,0}_{m,n+q}}{(m+1)(n+q+1)} - \frac{B^{0,0}_{m+p,n}}{(m+p+1)(n+1)} \\ &\qquad + \frac{B^{0,0}_{m+p,n+q}}{(m+p+1)(n+q+1)} = \sum_{m}^{m+p}\sum_{n}^{n+q} o\left(\frac{1}{m^2n^2}\right).\end{aligned}$$

Now, from Lemma 3 (iii) and H.-L., Lemmas 3 and 5 it follows that

$$\frac{B^{0,0}_{m,n+q}}{n+q+1} \quad \text{and} \quad \frac{B^{0,0}_{m+p,n}}{m+p+1}$$

are of the form $o(1/(n+q+1))$ and $o(1/(m+p+1))$ respectively, and therefore the second and third terms on the left-hand side of (31) tend to zero as p and q become infinite. Since the right-hand side of (31) also tends to a limit under the same conditions, it follows that the fourth term on the left-hand side of (31) does also; calling this limit h, and letting p and q become infinite, we have that

$$(32)\qquad \frac{B^{0,0}_{m,n}}{(m+1)(n+1)} = -h + o\left(\frac{1}{mn}\right).$$

Again, from (29),

$$(33)\quad b_{m+1,n+1} = -\frac{B^{0,0}_{m,n}}{(m+1)(n+1)} + \frac{B^{0,-1}_{m,n+1}}{m+1} + \frac{B^{-1,0}_{m+1,n}}{n+1} + o\left(\frac{1}{mn}\right).$$

Now by Lemmas 1 and 3 and H.-L., Lemmas 3 and 5 $B^{0,-1}_{m,n+1}$ is of the form $o(1/n)$, whence $B^{0,-1}_{m,n+1}/(m+1)$ is $o(1/mn)$. Similarly, $B^{-1,0}_{m+1,n}/(n+1)$ is $o(1/mn)$. Hence, from (32) and (33) we obtain (28) as desired.

(ii) We have next to prove the result in general case $r > 0$. We rewrite (21) as

$$(34)\qquad \begin{aligned}&(m+r+2)(n+r+2)B^{r,r}_{m,n}-(m+r+2)(n+1)B^{r,r}_{m,n+1}\\&-(m+1)(n+r+2)B^{r,r}_{m+1,n}+(m+1)(n+1)B^{r,r}_{m+1,n+1}\\&\qquad=A^{r,r}_{m,n}=o(m^r n^r).\end{aligned}$$

If we put

$$(35)\qquad B^{r,r}_{m,n}=(m+r+1)(m+r)\cdots(m+1)(n+r+1)(n+r)\cdots(n+1)\varphi_{m,n},$$

we obtain from (34)

$$(36)\qquad \begin{aligned}&\varphi_{m,n}-\varphi_{m+1,n}-\varphi_{m,n+1}+\varphi_{m+1,n+1}\\&=\frac{A^{r,r}_{m,n}}{(m+r+2)\cdots(m+1)(n+r+2)\cdots(n+1)}=o\left(\frac{1}{m^2n^2}\right),\end{aligned}$$

whence we deduce that

$$(37)\qquad \begin{aligned}&\varphi_{m,n}-\varphi_{m+p,n}-\varphi_{m,n+q}+\varphi_{m+p,n+q}\\&=\sum_{m}^{m+p}\sum_{n}^{n+q}\frac{A^{r,r}_{\mu,\nu}}{(\mu+r+2)\cdots(\mu+1)(\nu+r+2)\cdots(\nu+1)}\\&=\sum_{m}^{m+p}\sum_{n}^{n+q}o\left(\frac{1}{m^2n^2}\right).\end{aligned}$$

Again from Lemma 3 and H.-L., Lemmas 3 and 5, it follows that $\varphi_{m+p,n}$ and $\varphi_{m,n+q}$ are $o(1/(m+p))$ and $o(1/(n+q))$ respectively, and therefore approach zero as p and q become infinite. Hence we conclude that $\varphi_{m+p,n+q}$ approaches a limit, say $-\varphi$, as p and q become infinite. Therefore, letting p and q become infinite in (37), we have

$$(38)\qquad \begin{aligned}\varphi_{m,n}&=\varphi+o\left(\frac{1}{mn}\right)\\&=\varphi+\sum_{m}^{\infty}\sum_{n}^{\infty}\frac{A^{r,r}_{\mu,\nu}}{(\mu+r+2)\cdots(\mu+1)(\nu+r+2)\cdots(\nu+1)}.\end{aligned}$$

Thus (35) becomes

$$(39)\qquad B^{r,r}_{m,n}=(m+r+1)\cdots(m+1)\ (n+r+1)\cdots(n+1)\varphi+o(m^r n^r).$$

We are now in a position to deduce our lemma. We first put (19) and (20) in (21) to obtain

$$(40)\qquad \begin{aligned}A^{r,r}_{m,n}=o(m^r n^r)=&-(r+1)^2B^{r,r}_{m,n}+(m+1)(n+1)B^{r-1\,r-1}_{m+1,n+1}\\&+(r+1)[P^{r,r}_{m,n}+Q^{r,r}_{m,n}].\end{aligned}$$

We next put (19′) and (20′) in (40) to find

$$(41)\qquad \begin{aligned}A^{r,r}_{m,n}=o(m^r n^r)=&(r+1)^2B^{r,r}_{m,n}-(n+1)P^{r,r-1}_{m,n+1}\\&-(m+1)Q^{r-1,r}_{m+1,n}-(m+1)(n+1)B^{r-1,r-1}_{m+1,n+1}.\end{aligned}$$

Then, substituting (17) and (17′) in (41), we have

$$A^{r,r}_{m,n} = o(m^r n^r) = (r+1)^2 B^{r,r}_{m,n} - (m+1)(n+1) B^{r-1,r-1}_{m+1,n+1} + C^r_m o(n^r) + C^r_n o(m^r),$$

or

$$(42) \qquad (m+1)(n+1) B^{r-1,r-1}_{m+1,n+1} = (r+1)^2 B^{r,r}_{m,n} + o(m^r n^r).$$

Finally, if we put (39) into (42) and divide by $(m+1)(n+1)$, we obtain

$$B^{r-1,r-1}_{m+1,n+1} = h C^{r,r}_{m+1,n+1} + o(m^{r-1} n^{r-1}),$$

which is the desired result, (27), with $m = m+1$, $n = n+1$, and $h = [(r+1)!]^2 \varphi$. The lemma is then complete.

LEMMA 7. *If A, A_C, and A_R are summable (C, r) to zero, then* (i) *there are solutions of* (13′) *and* (14′) *such that Q_C and P_R are summable $(C, r-1)$ to zero, and* (ii) *there is a solution $b_{m,n}$ of* (13), (14) *and* (15) *such that B^* is summable $(C, r-1)$ to zero, where*

$$b^*_{m,n} = b_{m,n} - h^*,$$

h^* a constant.

Part (i) of the lemma follows immediately through an application of Lemmas 4 and 5 of H.-L. to the simple series involved.

To prove (ii), we argue that if $b_{m,n}$ is any solution of (13), (14) and (15), all others are of the form

$$b^*_{m,n} = b_{m,n} - h^*,$$

where h^* is a constant. If, in particular, $h^* = h$, we have

$$B^{*r-1,r-1}_{m,n} = B^{r-1,r-1}_{m,n} - h C^{r,r}_{m,n} = o(m^{r-1} n^{r-1}),$$

by Lemma 6, and hence B^* is summable $(C, r-1)$ to zero.

6. **Forms of the solutions of (13)–(15).** We now wish to determine the forms of the solutions mentioned in Lemma 7 as being necessary for the summability (C, r) of A, A_C, and A_R.

LEMMA 8. *If A, A_C, and A_R are summable (C, r) to zero, then* (i)

$$(43) \qquad A'' = \sum_0^\infty \sum_0^\infty \frac{a_{\mu,\nu}}{(\mu+1)(\nu+1)}$$

is summable $(C, r-1)$, where

$$(44) \quad A'' = [(r+1)!]^2 \sum_0^\infty \sum_0^\infty \frac{A^{r,r}_{\mu,\nu}}{(\mu+r+2)\cdots(\mu+1)(\nu+r+2)\cdots(\nu+1)};$$

and if

$$b_{m,n} = \sum_{m}^{\infty}\sum_{n}^{\infty} \frac{a_{\mu,\nu}}{(\mu+1)(\nu+1)}, \quad (C, r-1), \tag{45}$$

then B is summable $(C, r-1)$ to zero.

Also, (ii)

$$A' = \sum_{\nu=0}^{\infty} \frac{a_{m,\nu}}{\nu+1}, \text{ every } m, \tag{46}$$

is summable $(C, r-1)$, where

$$A' = (r+1)! \sum_{\nu=0}^{\infty} \frac{A_{m,\nu}^{-1,r}}{(\nu+r+2)\cdots(\nu+1)}, \quad \text{every } m; \tag{47}$$

and if

$$p_{m,n} = \sum_{\nu=n}^{\infty} \frac{a_{m,\nu}}{\nu+1}, \quad (C, r-1), \text{ every } m, \tag{48}$$

then P_R is summable $(C, r-1)$ to zero, as is, by Lemma 3 (ii), *B_R.*

(iii) *Similar results to those in* (ii) *hold for B_C, if we replace B_R by B_C, $p_{m,n}$ by $q_{m,n}$, and n by m.*

Parts (ii) and (iii) follow immediately from the analysis in our Lemma 7 and in Lemma 5 of H.-L.; in fact, they are in no wise dependent on our work, and were used freely in some of our previous deductions; they are included here merely for the sake of completeness.

For the proof of (i), we resume from Lemma 7 that B^* is summable $(C, r-1)$ to zero; thus

$$\sum_{0}^{\infty}\sum_{0}^{\infty} \{b^*_{m,n} - b^*_{m+1,n} - b^*_{m,n+1} + b^*_{m+1,n+1}\}$$

is summable $(C, r-1)$ to $b^*_{0,0}$, or A'' is summable $(C, r-1)$ to $b^*_{0,0}$. Thus there exists a $b_{m,n} = b^*_{m,n} + h^*$, which is plainly a solution of (13), (14) and (15); but $b_{m,n} \to 0$ $(C, r-1)$ as m and n become infinite, by definition, and $b^*_{m,n} \to 0$ $(C, r-1)$ since B^* is summable $(C, r-1)$; hence $h^* = 0$, and $b^*_{m,n} = b_{m,n}$.

We can now infer that, since h is zero, the φ of (38) is zero. Hence,

$$\begin{aligned} A'' &= b^*_{0,0} = b_{0,0} = B^{r,r}_{0,0} = [(r+1)!]^2 \varphi_{0,0} \\ &= [(r+1)!]^2 \sum_{0}^{\infty}\sum_{0}^{\infty} \frac{A^{r,r}_{\mu,\nu}}{(\mu+r+2)\cdots(\mu+1)(\nu+r+2)\cdots(\nu+1)}, \end{aligned}$$

by (35) and (38), and the lemma is complete.

7. Proof of Theorem I. We are now ready to prove Theorem I.

For the sufficiency of the conditions enunciated in the statement of the theorem we suppose that $A^{r+1,r+1}$, $A^{r+1,r}$ and $A^{r,r+1}$ are summable $(C, -1)$.

Then, by successive applications of Lemma 5, we conclude that $A^{\varrho,\varrho}$ and its rows and columns are summable $(C, r-\varrho)$, for $\varrho = r, r-1, r-2, \cdots, 0$; also, $A^{r,r} = A^{r-1,r-1} = \ldots = A^{\varrho,\varrho}$, with corresponding equalities for the sums of the simple series involved.

For the necessity of the conditions of the theorem we suppose that A, A_C, and A_R are summable (C, r). Then, by Lemmas 7 and 8 there exist numbers

$$a^{1,1}_{m,n} = b_{m,n} = \sum_{m}^{\infty}\sum_{n}^{\infty} \frac{a_{\mu,\nu}}{(\mu+1)(\nu+1)},$$

$$a^{1,0}_{m,n} = q_{m,n} = \sum_{\mu=m}^{\infty} \frac{a_{\mu,n}}{\mu+1}, \quad \text{(every } n\text{)},$$

$$a^{0,1}_{m,n} = p_{m,n} = \sum_{\nu=n}^{\infty} \frac{a_{m,\nu}}{\nu+1}, \quad \text{(every } m\text{)},$$

such that $A^{1,1}$, $A^{1,0}$, and $A^{0,1}$, and hence $A^{1,1}$ and its rows and columns, are summable $(C, r-1)$. Successive applications of this reasoning lead to the conditions as outlined. Theorem I is therefore established.

8. **Theorem II.** For application to the double Fourier's series, to be given in a later paper, we shall need a slightly different form of Theorem I, the difference amounting to little more than a change in notation. For this purpose we supplant the equations in the statement of Theorem I by new ones in which the factors $m+1$ and $n+1$ of (4), (5) and (6) are changed to m and n, and the factors $1/(\mu+1)$ and $1/(\nu+1)$ of (10), (11) and (12) are changed to $1/\mu$ and $1/\nu$; moreover, we provide that any term of the new series which contains a zero subscript shall be zero. We thus obtain a new series which differs from the other one only in the addition of a row and a column of zeroes, which does not affect the summability or the sum of the series. Thus we have Theorem II:

THEOREM II. *The results of Theorem* I *are valid under the replacements noted above.*

THE UNIVERSITY OF CINCINNATI,
HARVARD UNIVERSITY.

CERTAIN CONFIGURATIONS ON CUBICS.*

By Frank Chappell Ogg.

1. **Introduction.** The object of this paper is to investigate the configurations existing among the in-circumscribed polygons on a general plane cubic curve, with particular reference to their group properties. Triangles and quadrangles will be studied in detail, with some attention to p-gons (p an odd prime) and to systems of cubics, each member of which has the same special modulus.

Porter† has discussed by means of elliptic functions certain relations among the in-circumscribed triangles. The application of elliptic functions to this and related problems has been considered by Emch‡ and Halphen.§ The latter has also showed the existence of in-circumscribed triangles by a method not involving elliptic functions. Disteli‖ has discussed the problem of closing polygons on a cubic from the view point of pure geometry.

Kantor¶ has treated by synthetic methods a certain 1 : 3 mapping of a plane upon a surface of class and order 3 and has obtained certain configurations among the in-circumscribed triangles. Many of the results are given without proof.

2. **The threefold mapping of a plane upon a cubic surface.** If we consider the fundamental triangle $A_1(1,0,0)$, $A_2(0,1,0)$, $A_3(0,0,1)$, and form the triangular polar of a point (x_1, x_2, x_3) with respect to $A_1 A_2 A_3$, we find it to be, in variables (y), the line

$$L_1\colon\ y_1 x_2 x_3 + y_2 x_3 x_1 + y_3 x_1 x_2 = 0.$$

Similarly the triangular polars of $(\omega x_1, \omega^2 x_2, x_3)$ and $(\omega^2 x_1, \omega x_2, x_3)$, where $\omega^3 = 1$, are respectively

* Received September 31, 1927.

† Porter, M. B., Sets of coincidence points on the non-singular cubics of syzygetic sheaf, Trans. Amer. Math. Soc., vol. 2 (1901), pp. 37–42.

‡ Emch, Arnold, Application of elliptic functions to problems of closure, Univ. of Colorado Studies, vol. 1, no. 2 (1902), pp. 81–133.

§ Halphen, G. H., Sur les courbes du troisième degré, Math. Ann., vol. 15 (1879), pp. 359–379.

‖ Disteli, M., Die Steinerschen Schließungsprobleme nach darstellender Methode.

¶ Kantor, S., Über eine ein-dreideutige ebene Abbildung einer Fläche dritter Ordnung, Journ. für Math., vol. 95 (1883), pp. 147–163.

$$L_2\colon\ y_1\omega^2 x_2x_3+y_2\omega x_3x_1+y_3x_1x_2=0,$$
$$L_3\colon\ y_1\omega x_2x_3+y_2\omega^2 x_3x_1+y_3x_1x_2=0.$$

Any pair of the lines L_i meet in the point whose triangular polar is the third line. Thus the trilateral $L_1L_2L_3$ is self polar as to $A_1A_2A_3$.

The plane (x) may be mapped $3:1$ upon a surface by means of the four linearly independent cubics on the three points as follows:

$$(1)\qquad \begin{aligned} v_1 &= -x_1^3+x_2^3+x_3^3, & v_2 &= x_1^3-x_2^3+x_3^3,\\ v_3 &= x_1^3+x_2^3-x_3^3, & v_4 &= 2x_1x_2x_3. \end{aligned}$$

The surface obtained by eliminating the x_i is

$$(2)\qquad (v_1+v_2)(v_2+v_3)(v_3+v_1)-v_4^3=0.$$

The cubic surface (2) is cut by the plane $v_4=0$ in three lines. To the pencil of planes α on a generic line s in the plane $v_4=0$ corresponds in (x) a syzygetic pencil of cubics, for the three points in which s cuts the three lines

$$v_1+v_2=0,\ v_4=0;\quad v_2+v_3=0,\ v_4=0;\quad v_3+v_1=0,\ v_4=0,$$

correspond nine points in (x) and to these lines themselves correspond the triply covered lines

$$x_3^3=0,\quad x_1^3=0,\quad x_2^3=0.$$

Thus the nine points are flexes and $A_1A_2A_3$ is a flex trilateral.

The surface (2) is cut by a generic plane

$$(3)\qquad a_1v_1+a_2v_2+a_3v_3+a_4v_4=0,$$

containing s, in a plane cubic. To this cubic corresponds in (x) the cubic

$$(4)\quad x_1^3(-a_1+a_2+a_3)+x_2^3(a_1-a_2+a_3)+x_3^3(a_1+a_2-a_3)+2a_4x_1x_2x_3=0,$$

or

$$b_1x_1^3+b_2x_2^3+b_3x_3^3+2a_4x_1x_2x_3=0.$$

To any point of the cubic in (3) correspond three points of (4), which are a cycle of the collineation:

$$\varrho x_1'=\omega x_1,\quad \varrho x_2'=\omega^2x_2,\quad \varrho x_3'=x_3.$$

To three of the nine flexes of the cubic in (3), namely those in s, correspond A_1, A_2, A_3. To each of the other six flexes correspond three points

on (4). The tangent plane β to (2) at one of these flexes cuts (2) in a cubic with a double point. This cubic maps in (x) into a cubic with three double points, a triangle. One nodal tangent of the cubic in β coincides with the flex tangent of the cubic in (3); the other is distinct. Thus the correspondents of (3) and of β meet in sets of two coincident points and a distinct point, that is to say, an in-circumscribed triangle of (4). Thus the existence of six in-circumscribed tringles is established.

In order to establish the existence of additional in-circumscribed triangles we take on the cubic in (3) any McLaurin line s' not belonging to the same flex trilateral as s. Pass any plane α' through s'. α' cuts (2) in a general cubic. Let α' have the equation $v_4' = 0$. In α' take any triangle A_1', A_2', A_3' containing on each side one flex of the cubic in (3). The line $A_i' A_j'$ and the flex tangent intersecting it determine a plane. For A_4' take any point not on α' nor the planes just determined. We find the fourth harmonic of the plane on $A_i' A_j'$ with respect to $v_4' = 0$ and $v_k' = 0$, i, j, k being 1, 2, 3 in any order. These three fourth harmonic planes intersect in a point E which may be chosen as (1, 1, 1, 1). Then the planes determined by the lines $A_i' A_j'$ and the intersecting flex tangents will have the equations

$$v_1' + v_2' = 0, \quad v_2' + v_3' = 0, \quad v_3' + v_1' = 0.$$

Thus we have a pencil of cubic surfaces

$$(v_1' + v_2')(v_2' + v_3')(v_3' + v_1') - \lambda v_4'^3 = 0.$$

The parameter λ may be so adjusted that the above surface shall contain the cubic C in (3), which may be mapped by this surface into the same cubic in (x) as in the mapping by (2). Thus we have established a sextuple of in-circumscribed triangles for each flex trilateral of a plane cubic lying on (2). Thus there are twenty-four triangles containing seventy-two points.

It is desired to determine the condition exerted on planes cutting (2) such that the cubics in the resulting set of images in (x) shall be harmonic or equianharmonic. We shall consider the case $a_4 = 0$ separately since it is obvious that this leads to a net of equianharmonic cubics. We here take (4) in the second form with $a_4 = \frac{1}{2}$. The Hessian of (4) is

$$(5) \qquad H \equiv \begin{vmatrix} 6b_1 x_1 & x_3 & x_2 \\ x_3 & 6b_2 x_2 & x_1 \\ x_2 & x_1 & 6b_3 x_3 \end{vmatrix} = 0,$$

or

$$b_1 x_1^3 + b_2 x_2^3 + b_3 x_3^3 - M x_1 x_2 x_3 = 0; \qquad M = \frac{1 + 108 b_1 b_2 b_3}{3}.$$

The Hessian of H is

$$H' \equiv \begin{vmatrix} 6b_1x_1 & -Mx_3 & -Mx_2 \\ -Mx_3 & 6b_2x_2 & -Mx_1 \\ -Mx_2 & -Mx_1 & 6b_3x_3 \end{vmatrix} = 0, \tag{6}$$

or

$$b_1x_1^3 + b_2x_2^3 + b_3x_3^3 - \frac{108b_1b_2b_3 - M^3}{3M^2}x_1x_2x_3 = 0.$$

If (4) is harmonic it must be identical with the Hessian of its Hessian, or (4) is the same as (6) and we have

$$(108b_1b_2b_3)^3 - 6(108b_1b_2b_3)^2 - 42(108b_1b_2b_3) - 8 = 0. \tag{7}$$

If (4) is equianharmonic, H' will coincide with H. Then we have

$$4(108b_1b_2b_3)^3 + 12(108b_1b_2b_3)^2 - 15(108b_1b_2b_3) + 4 = 0. \tag{8}$$

Thus we have the

THEOREM. *The condition that the images in* (x) *of the cubics cut out on* (2) *by the system of planes*

$$a_1v_1 + a_2v_2 + a_3v_3 + \frac{1}{2}v_4 = 0$$

shall be harmonic or equianharmonic is of order nine in either case.

We note that both (7) and (8) are cubics in the unknown $b_1b_2b_3$; hence each has three roots $b_1b_2b_3 = r_i$, $i = 1, 2, 3$. If we now let

$$a_1 = u_1/u_4, \quad a_2 = u_2/u_4, \quad a_3 = u_3/u_4,$$

we have

$$(-u_1 + u_2 + u_3)(u_1 - u_2 + u_3)(u_1 + u_2 - u_3) - r_iu_4^3 = 0, \tag{9}$$

a surface of class three.

Thus there are in either case three systems of cubics, corresponding to three class surfaces, with the required special modulus. In (7) the three roots r_i are distinct but in (8) two of the roots are equal. Thus there are, in effect, three systems of harmonic but only two distinct systems of equianharmonic cubics, if we omit the set due to $a_4 = 0$.

3. The in-circumscribed triangles of a plane elliptic cubic. The vertices of in-circumscribed triangles are points with the elliptic parameter u such that $9u \equiv 0 \pmod{2\omega_1, 2\omega_2}$, or

$$u = \frac{2m_1\omega_1 + 2m_2\omega_2}{9}.$$

Letting m_1, m_2 vary independently from 0 to 8 we get 81 values, from which must be deducted the nine where $3u \equiv 0$ (the flexes). Thus there are 72 proper coincidence points or 24 triangles. Porter showed that if two in-circumscribed triangles are perspective from a flex, they are perspective in three ways, that is, from three flexes. In this section we will discuss the permutation of the vertices and of the triangles under all the operations of Γ_{18}, the group under which a general cubic is invariant.

It is well known that Γ_{18} may be generated by three involutory perspectives, not belonging to the same subgroup of order six. If the flexes are denoted by the letters $a, b, c, d, e, f, g, h, k$ these three involutions may be taken as

(T) $\qquad bc \cdot dg \cdot ek \cdot fh,$

(U) $\qquad ac \cdot dk \cdot eh \cdot fg,$

(V) $\qquad ag \cdot bk \cdot ch \cdot ef.$

In order to fix the notation the 72 points were labeled $1, 2, \cdots, 36$ and $1', 2', \cdots, 36'$, such that T sends j into j'. U may be written in terms of the vertices:

(1 2′) (2 3′) (3 1′) (4 5′) (5 6′) (6 4′) (7 8′) (8 9′) (9 7′) (10 17′) (11 18′) (12 16′) (13 26′) (14 27′) (15 25′) (16 23′) (17 24′) (18 22′) (19 14′) (20 15′) (21 13′) (22 11′) (23 12′) (24 10′) (25 20′) (26 21′) (27 19′) (28 36′) (29 34′) (30 35′) (31 28′) (32 29′) (33 30′) (34 32′) (35 33′) (36 31′).

V may be written:

(1 4′) (2 5′) (3 6′) (4 7′) (5 8′) (6 9′) (7 1′) (8 2′) (9 3′) (10 16′) (11 17′) (12 18′) (13 19′) (14 20′) (15 21′) (16 22′) (17 23′) (18 24′) (19 25′) (20 26′) (21 27′) (22 10′) (23 11′) (24 12′) (25 13′) (26 14′) (27 15′) (28 30′) (29 28′) (30 29′) (31 33′) (32 31′) (33 32′) (34 36′) (35 34′) (36 35′).

The eight cyclic collineations of the third order each leave the triangles of one sextuple invariant (permuting the vertices cyclically) and permute the triangles of the other sextuples in sets of three; this may be illustrated by

$U \times T =$ (1 2 3) (4 5 6) (7 8 9) (10 17 24) (11 18 22) (12 16 23) (13 26 21) (14 27 19) (15 25 20) (28 36 31) (29 34 32) (30 35 33) (1′ 3′ 2′) (4′ 6′ 5′) (7′ 9′ 8′) (10′ 24′ 17′) (11′ 22′ 18′) (12′ 23′ 16) (13′ 21′ 26′) (14′ 19′ 27′) (15′ 20′ 25′) (28′ 31′ 36′) (29′ 32′ 34′) (30′ 33′ 35′).

It may readily be verified that the 72 points or 24 triangles divide into sets, as follows:

1. Two sets Σ and Σ' such that any point of Σ is perspective with nine points of Σ' (from the nine flexes). These nine points are the vertices of three triangles.

2. Four sextuples of triangles

$$\Delta_i, \quad i = 1, 2, 3, 4,$$

each containing three accented triangles and three non-accented ones, such that under Γ_{18} triangles of Δ_i go into triangles of the same Δ_i.

3. Γ_{18} contains as a subgroup the symmetric group on x_1, x_2, x_3 of order 6.* The six points obtained by the permutation of the coördinates of one lie on a conic; whence the 72 points lie by sixes on 12 conics. The product of these conics and the eighth power of the cubic form two curves of a pencil of curves of order 24, invariant under the group.

4. **In-circumscribed quadrilaterals on C_3.** As in the case of triangles it may be shown that the vertices u of in-circumscribed quadrilaterals will be such that

$$15u \equiv 0 \pmod{2\omega_1, 2\omega_2}.$$

Deducting the flexes, we have $15^2 - 9 = 216$ vertices or 54 quadrilaterals. The 54 quadrilaterals are divided into nine sets of six each, a set being uniquely associated with a flex. The sets may be named $A_1, A_2, \cdots, A_{24}$; $B_1, B_2, \cdots, B_{24}; \cdots; K_1, K_2, \cdots, K_{24}$. The groups of subscripts 1–4, 5–8, $\cdots$, 21–24 denote quadruples of points which belong to the same quadrilateral.

The relations of these points under $T = bc \cdot dg \cdot ek \cdot fh$ are:

$$A_1A_3 \cdot A_2A_4 \cdot A_5A_7 \cdot A_6A_8 \cdot A_9A_{11} \cdot A_{10}A_{12} \cdot A_{13}A_{15} \cdot A_{14}A_{16} \cdot$$
$$A_{17}A_{19} \cdot A_{18}A_{20} \cdot A_{21}A_{23} \cdot A_{22}A_{24} \cdot$$
$$B_iC_i \cdot D_iG_i \cdot E_iK_i \cdot F_iH_i, \quad i = 1, 2, \cdots, 24.$$

$U = ac \cdot dk \cdot eh \cdot fg$ gives:

$$B_1B_3 \cdot B_2B_4 \cdot B_5B_7 \cdot B_6B_8 \cdot B_9B_{11} \cdot B_{10}B_{12} \cdot B_{13}B_{15} \cdot B_{14}B_{16} \cdot$$
$$B_{17}B_{19} \cdot B_{18}B_{20} \cdot B_{21}B_{23} \cdot B_{22}B_{24} \cdot$$
$$A_iC_i \cdot D_iK_i \cdot E_iH_i \cdot F_iG_i, \quad i = 1, 2, \cdots, 24.$$

Similarly V (center d) holds the "D" quadrilaterals fixed, interchanging their opposite corners, and permutes the rest with their associated flexes.

As in the case of triangles, the vertices lie on conics by sixes. The number of conics is 36 and their product with C_2^{24} generate a pencil of invariant curves of order 72.

5. **Steiner quadruples.** The Steiner quadruple associated with a point whose elliptic parameter is u is given by

$$-\frac{u}{2}, \quad -\frac{u}{2}+\omega_1, \quad -\frac{u}{2}+\omega_2, \quad -\frac{u}{2}+\omega_1+\omega_2$$

* Γ_{18} contains four subgroups of order 6; the one referred to here may be written as a substitution group on the three *real* flexes, assumed to be a, b, c.

where $2\omega_1$, $2\omega_2$ are periods. The Steiner quadruple associated with a coincidence point will clearly consist of another coincidence point and three new points. The 216 new points, in the case of triangles, will be distinct.

It is well known that if the points of one quadruple are joined to those of another, the 16 lines cut the cubic in four points of a Steiner quadruple. If we consider the quadruples associated with vertices of in-circumscribed triangles, it appears that if the quadruples attached to a vertex and its tangential are joined, the 16 lines reduce to 10 and the first quadruple is reproduced. If we choose two vertices which are perspective from a flex and join the points of the quadruples, the lines cut out the flex and the three intersections of the cubic with its satellite.

The same situation is disclosed with respect to the vertices of quadrilaterals.

6. **In-circumscribed p-gons.** The number of points which are vertices of in-circumscribed p-gons (p an odd prime) may be determined by the same method as for triangles. The number is

$$(2^p+1)^2-9 = 2^3(2^{p-1}-1)(2^{p-2}+1).$$

This number is divisible by p, since $2^{p-1} \equiv 1, \bmod p$.

In no case can a polygon, the number of whose sides is an odd prime, be invariant under a perspective from a flex, for, if so, an even number of vertices would be interchanged in pairs and the remaining odd number would be invariant points. But the invariant points on C_3 under perspective from a flex are the flex itself and its three tangentials, one of which is the vertex of an in-circumscribed polygon. Thus from every flex a p-gon is projected into a distinct p-gon. As an example we consider the two polygons perspective from (a):

$$\begin{aligned}
P_1 &= u = \frac{2m_1\omega_1 + 2m_2\omega_2}{2^p+1},\\
P_2 &= -2u,\\
P_3 &= 4u,\\
&\cdots\\
P_{p-1} &= (-2)^{p-2}u,\\
P_p &= (-2)^{p-1}u.
\end{aligned}$$

The vertices of P' perspective to P from (a) are:

$$\begin{aligned}
P'_1 &= -u,\\
P'_2 &= 2u,\\
P'_3 &= -4u,\\
&\cdots\\
P'_{p-1} &= -(-2)^{p-2}u,\\
P'_p &= -(-2)^{p-1}u.
\end{aligned}$$

In discussing triangles we observed that if two triangles are perspective at all, they are perspective in three ways. But in the case of p-gons, $p>3$, no such situation is found. Indeed, if we join P_i to P'_{i+k}, reducing the subscripts modulo p, the residual intersections of these joins with C_3 form the vertices of a new p-gon. Thus we have the

THEOREM. *If two p-gons P and P' are perspective from a flex and the vertices of P are joined to the vertices of P', advancing the subscripts for P' cyclically, the residual intersections of these joins with C_3 are the vertices of a new p-gon. When $p=3$, the new p-gon collapses into a flex tangent, the three flexes obtained as centers being collinear.*

7. The pencil of cubics. $f \equiv x_1^2x_2+x_2^2x_3+x_3^2x_1+\lambda(x_1^2x_3+x_2^2x_1+x_3^2x_2)=0$.

The two given cubics of the pencil are equianharmonic and have three of their nine common points at the vertices of the fundamental triangle. This triangle is in-circumscribed to each, the relative order of the points being opposite for the two curves. The other six base points are found to be

$$(\omega, 1, 1) = B_1, \quad (1, \omega, 1) = B_2, \quad (1, 1, \omega) = B_3,$$
$$(\omega^2, 1, 1) = C_1, \quad (1, \omega^2, 1) = C_2, \quad (1, 1, \omega^2) = C_3.$$

B_i, C_i are found to be arranged in two in-circumscribed triangles to the two given curves, the relative order of the vertices being opposite for the two curves as for A_i.

The base points of f form three cycles of each of three cyclic perspectives of order three, such that the center of one is the intersection of the axes of the other two. The three perspectives are

$$1.\quad \begin{aligned} \rho x_1' &= \omega x_1+x_2+x_3,\\ \rho x_2' &= x_1+\omega x_2+x_3,\\ \rho x_3' &= x_1+x_2+\omega x_3, \end{aligned}$$

with center $(1, 1, 1)$ and axis $x_1+x_2+x_3=0$;

$$2.\quad \begin{aligned} \rho x_1' &= x_1+\omega x_2+x_3,\\ \rho x_2' &= x_1+x_2+\omega x_3,\\ \rho x_3' &= \omega x_1+x_2+x_3, \end{aligned}$$

with center $(\omega, \omega^2, 1)$ and axis $x_1+\omega^2 x_2+\omega x_3=0$;

$$3.\quad \begin{aligned} \rho x_1' &= x_1+x_2+\omega x_3,\\ \rho x_2' &= \omega x_1+x_2+x_3,\\ \rho x_3' &= x_1+\omega x_2+x_3, \end{aligned}$$

with center $(\omega^2, \omega, 1)$ and axis $x_1+\omega x_2+\omega^2 x_3=0$.

It may at once be verified that the nine points A_i, B_i, C_i, $i = 1, 2, 3$, lie by threes on nine lines which are concurrent by threes at $(1, 1, 1)$, $(\omega, \omega^2, 1)$, and $(\omega^2, \omega, 1)$. The three triples of lines form three degenerate cubics of f and correspond to $\lambda = -1, -\omega, -\omega^2$. These are the only degenerate cubics in f.

The Hessian of a curve of f is

$$H = \begin{vmatrix} x_2+\lambda x_3 & x_1+\lambda x_2 & x_3+\lambda x_1 \\ x_1+\lambda x_2 & x_3+\lambda x_1 & x_2+\lambda x_3 \\ x_3+\lambda x_1 & x_2+\lambda x_3 & x_1+\lambda x_2 \end{vmatrix} = 0,$$

which reduces to

$$(1+\lambda^3)(x_1+x_2+x_3)(x_1+\omega x_2+\omega^2 x_3)(x_1+\omega^2 x_2+\omega x_3) = 0.$$

For $1+\lambda^3 \neq 0$, H is independent of λ; that is to say, all the proper cubics of the pencil have the same Hessian, the above trilateral. Thus we have the

THEOREM. *Every proper cubic of the pencil f is equianharmonic.*

It is known that the flex-tangents at three collinear flexes of an equianharmonic cubic are concurrent. It is readily verified that the tangents at the three flexes of a curve of f on a side of H are concurrent at the opposite vertex of H. Indeed, the Cayleyan of every curve of f consists of these three points $(1, 1, 1)$, $(\omega, \omega^2, 1)$, $(\omega^2, \omega, 1)$.

The flexes of curves of f are variable with λ. Let

$$\alpha = \left(\frac{\lambda+\omega}{\lambda\omega+1}\right)^{1/3}, \qquad \beta = \left(\frac{\lambda\omega^2+\omega}{\lambda+1}\right)^{1/3}, \qquad \gamma = \left(\frac{\lambda\omega+\omega^2}{\lambda+1}\right)^{1/3}.$$

The coördinates of the flexes are found to be:

$$\begin{aligned}
&R_1\colon (\omega^2-\omega\alpha,\ \omega-\omega^2\alpha,\ 1-\alpha),\\
&R_2\colon (\omega^2-\omega^2\alpha,\ \omega-\alpha,\ 1-\omega\alpha),\\
&R_3\colon (\omega^2-\alpha,\ \omega-\omega\alpha,\ 1-\omega^2\alpha),\\
&S_1\colon (\omega-\beta,\ \omega^2-\beta,\ 1-\beta),\\
&S_2\colon (\omega-\omega\beta,\ \omega^2-\omega\beta,\ 1-\omega\beta),\\
&S_3\colon (\omega-\omega^2\beta,\ \omega^2-\omega^2\beta,\ 1-\omega^2\beta),\\
&T_1\colon (\omega^2-\gamma,\ \omega-\gamma,\ 1-\gamma),\\
&T_2\colon (\omega^2-\omega^2\gamma,\ \omega-\omega^2\gamma,\ 1-\omega^2\gamma),\\
&T_3\colon (\omega^2-\omega\gamma,\ \omega-\omega\gamma,\ 1-\omega\gamma).
\end{aligned}$$

The flexes lie by threes on twelve lines, forming four trilaterals, as follows:

$$\begin{aligned} &R_1R_2R_3,\ S_1S_2S_3,\ T_1T_2T_3, && \text{fixed } (H),\\ &R_1S_1T_1,\ R_2S_2T_2,\ R_3S_3T_3, && \text{variable},\\ &R_1S_2T_3,\ R_2S_3T_1,\ R_3S_1T_2, && \text{variable},\\ &R_1S_3T_2,\ R_2S_1T_3,\ R_3S_2T_1, && \text{variable}. \end{aligned}$$

It is desired to obtain the locus enveloped by these nine variable lines as λ varies. It appears that all nine envelop the same locus, but, to obtain it, it is more convenient to transform f into

$$f' \equiv x_1^3+x_2^3+x_3^3+\varkappa(a_1x_1^3+a_2x_2^3+a_3x_3^3) = 0.$$

The Hessian of f' is H':

$$H' \equiv 216(1+\varkappa a_1)(1+\varkappa a_2)(1+\varkappa a_3)x_1x_2x_3 = 0.$$

It is clear that proper cubics of this pencil have a common Hessian, the fixed trilateral $x_1=x_2=x_3=0$. If we let $a_1-a_2=b_3$, $a_2-a_3=b_1$, $a_3-a_1=b_2$, the base points of f' are seen to be

$$\begin{aligned} &(b_1^{1/3}, b_2^{1/3}, b_3^{1/3}), && (\omega b_1^{1/3}, b_2^{1/3}, b_3^{1/3}), && (\omega^2 b_1^{1/3}, b_2^{1/3}, b_3^{1/3}),\\ &(b_1^{1/3}, \omega b_2^{1/3}, b_3^{1/3}), && (b_1^{1/3}, \omega^2 b_2^{1/3}, b_3^{1/3}), && (b_1^{1/3}, b_2^{1/3}, \omega b_3^{1/3}),\\ &(b_1^{1/3}, b_2^{1/3}, \omega^2 b_3^{1/3}), && (\omega b_1^{1/3}, \omega^2 b_2^{1/3}, b_3^{1/3}), && (\omega^2 b_1^{1/3}, \omega b_2^{1/3}, b_3^{1/3}). \end{aligned}$$

If f' is written in variables x_i', it may be carried into f by

$$(T) \qquad \begin{aligned} \rho x_1' &= r(x_1+x_2+x_3),\\ \rho x_2' &= s(x_1+\omega x_2+\omega^2 x_3),\\ \rho x_3' &= t(x_1+\omega^2 x_2+\omega x_3), \end{aligned}$$

where

$$r = \left(\frac{\lambda+1}{1+\varkappa a_1}\right)^{1/3}, \qquad s = \left(\frac{\lambda\omega+\omega^2}{1+\varkappa a_2}\right)^{1/3}, \qquad t = \left(\frac{\lambda\omega^2+\omega}{1+\varkappa a_3}\right)^{1/3}$$

where r, s, t are clearly not zero for values of λ and $\varkappa$ corresponding to proper cubics.

The inverse of T is

$$(T^{-1}) \qquad \begin{aligned} \sigma x_1 &= \frac{x_1'}{r}+\frac{x_2'}{s}+\frac{x_3'}{t},\\ \sigma x_2 &= \frac{x_1'}{r}+\frac{\omega^2 x_2'}{s}+\frac{\omega x_3'}{t},\\ \sigma x_3 &= \frac{x_1'}{r}+\frac{\omega x_2'}{s}+\frac{\omega^2 x_3'}{t}. \end{aligned}$$

In f' as in f the variable McLaurin lines envelop a locus. One of the lines is

$$r_1^{1/3} x_1 + r_2^{1/3} x_2 + r_3^{1/3} x_3 = 0,$$

where $r_i = 1 + \varkappa a_i$. The tangential equation of the locus is

$$b_1 u_1^3 + b_2 u_2^3 + b_3 u_3^3 = 0.$$

The point equation is

$$b_2^2 b_3^2 x_1^6 + b_3^2 b_1^2 x_2^6 + b_1^2 b_2^2 x_3^6 - 2 b_1^2 b_2 b_3 x_2^3 x_3^3 - 2 b_1 b_2^2 b_3 x_3^3 x_1^3 - 2 b_1 b_2 b_3^2 x_1^3 x_2^3 = 0.$$

This locus has nine cusps whose coördinates and cuspidal tangents are as follows:

$$\begin{aligned}
&A\colon (0,\ b_2^{1/3},\ b_3^{1/3}), && b_3^{1/3} x_2 - b_2^{1/3} x_3 = 0.\\
&B\colon (0,\ \omega^2 b_2^{1/3},\ b_3^{1/3}), && \omega b_3^{1/3} x_2 - b_2^{1/3} x_3 = 0.\\
&C\colon (0,\ \omega b_2^{1/3},\ b_3^{1/2}), && \omega^2 b_3^{1/3} x_2 - b_2^{1/3} x_3 = 0.\\
&D\colon (b_1^{1/3},\ 0,\ b_3^{1/3}), && b_3^{1/3} x_1 - b_1^{1/3} x_3 = 0.\\
&E\colon (\omega b_1^{1/3},\ 0,\ b_3^{1/3}), && \omega^2 b_3^{1/3} x_1 - b_1^{1/3} x_3 = 0.\\
&F\colon (\omega^2 b_1^{1/3},\ 0,\ b_3^{1/3}), && \omega b_3^{1/3} x_1 - b_1^{1/3} x_3 = 0.\\
&G\colon (b_1^{1/3},\ b_2^{1/3},\ 0), && b_2^{1/3} x_1 - b_1^{1/3} x_2 = 0.\\
&H\colon (\omega^2 b_1^{1/3},\ b_2^{1/3},\ 0), && \omega b_2^{1/3} x_1 - b_1^{1/3} x_2 = 0.\\
&K\colon (\omega b_1^{1/3},\ b_2^{1/3},\ 0), && \omega^2 b_2^{1/3} x_1 - b_1^{1/3} x_2 = 0.
\end{aligned}$$

A, B, C lie on $x_1 = 0$, D, E, F on $x_2 = 0$, G, H, K on $x_3 = 0$; but the cusps do not lie by threes on any other lines, for general values of b_i. Whence the

THEOREM: *For a pencil of the type f' the nine variable McLaurin lines envelop a locus of order 6 and class 3, with nine cusps which lie on the Hessian of f'. The cuspidal tangents intersect by threes in the base points of f', and each cuspidal tangent passes through three base points. The tangents at the cusps which lie in the fundamental triangle intersect on the opposite vertex.*

8. **The cubic curve.** $C \equiv x_1^2 x_2 + x_2^2 x_3 + x_3^2 x_1 + x_1^2 x_3 + x_2^2 x_1 + x_3^2 x_2 = 0$.

This curve is invariant under a group $\overline{\Gamma}_{54}$, since it is equianharmonic, and also under the quadratic involution

$$Q\colon\ \rho x_i' = 1/x_i,\quad i = 1, 2, 3.$$

It is desired to obtain the group $\overline{\Gamma}_{54}$ and augment it by Q. C is the only proper cubic of f which is invariant under Q, the other cubics being

paired; the cubic belonging to the parameter λ_0 going into that belonging to $1/\lambda_0$.

A collineation carrying C into $N \equiv x_1^3+x_2^3+x_3^3 = 0$ is

$$(T) \qquad \begin{aligned} \varrho x_1 &= -2^{1/3}(x_1'+x_2'+x_3'),\\ \varrho x_2 &= x_1'+\omega x_2'+\omega^2 x_3',\\ \varrho x_3 &= x_1'+\omega^2 x_2'+\omega x_3'. \end{aligned}$$

The inverse is

$$(T^{-1}) \qquad \begin{aligned} \sigma x_1' &= 2^{-1/3}x_1 - x_2 - x_3,\\ \sigma x_2' &= 2^{-1/3}x_1 - \omega^2 x_2 - \omega x_3,\\ \sigma x_3' &= 2^{-1/3}x_1' - \omega x_2 - \omega^2 x_3. \end{aligned}$$

Γ_{54}, belonging to N, is generated by

$$cfk\cdot bhe = (\omega x_1, x_2, x_3), \qquad cfk\cdot agd = (x_1, \omega^2 x_2, x_3),$$
$$ace\cdot dfh\cdot bgk = (x_2, x_3, \omega x_1), \qquad ag\cdot bk\cdot ch\cdot ef = (x_1, \omega x_3, \omega^2 x_2).$$

$\bar{\Gamma}_{54} = T\Gamma_{54}T^{-1}$ is then generated by*

$$\begin{aligned} \varrho x_1' &= \omega^2 x_1+x_2+x_3, & \varrho x_1' &= x_1+x_2+x\omega_3,\\ \varrho x_2' &= x_1+\omega^2 x_2+x_3, & \varrho x_2' &= \omega x_1+x_2+x_3,\\ \varrho x_3' &= x_1+x_2+\omega^2 x_3; & \varrho x_3' &= x_1+\omega x_2+x_3; \end{aligned}$$

$$\begin{aligned} \varrho x_1' &= \omega\varepsilon x_1+\omega\delta x_2+\omega\theta x_3, & \varrho x_1' &= x_2,\\ \varrho x_2' &= \omega^2\theta x_1+\omega^2\varepsilon x_2+\omega^2\delta x_3, & \varrho x_2' &= x_1,\\ \varrho x_3' &= \delta x_1+\theta x_2+\varepsilon x_3; & \varrho x_3' &= x_3. \end{aligned}$$

The group generated by the above substitutions and Q is of order 108. Whence the

THEOREM. *Every equianharmonic cubic is invariant under a Cremona group of order* 108.

9. Certain types of bases for a pencil of cubics.

(a) Basis consisting of three cycles under a cyclic perspective of order three.

We select the cyclic perspective $R = (\omega x_1, x_2, x_3)$, which has the center (1, 0, 0) and axis $x_1 = 0$. We choose three lines through the center carrying the three cycles as $x_2 = 0$, $x_3 = 0$, $x_2 - x_3 = 0$, as we may do without loss of generality. Three points are selected, one on each line, aud linearly independent. They with their correspondents under R and R^2 form the basis. Let the three points be $(\xi, 0, 1)$, $(1, 1, 1)$, $(\eta, 1, 0)$. The

* $\varepsilon = -\omega^2+2^{1/3}\omega^2-4^{1/3}$, $\delta = -1+2^{1/3}\omega-4^{1/3}$, $\theta = -\omega+2^{1/3}-4^{1/3}$.

other points are then $(\omega\xi, 0, 1)$, $(\omega^2\xi, 0, 1)$, $(\omega, 1, 1)$, $(\omega^2, 1, 1)$, $(\omega\eta, 1, 0)$, $(\omega^2\eta, 1, 0)$. Passing a general cubic through these nine points we find that it belongs to the system

$$x_1^3 - \eta^3 x_2^3 - \xi^3 x_3^3 + e x_2^2 x_3 + k x_2 x_3^2 = 0.$$

The Hessian of a curve of this system is

$$\begin{vmatrix} 6x_1 & 0 & 0 \\ 0 & -6\eta^3 x_2 + 2ex_3 & 2ex_2 + 2kx_3 \\ 0 & 2ex_2 + 2kx_3 & -6\xi^3 x_3 + 2kx_2 \end{vmatrix} = 0,$$

which consists on three non-concurrent lines unless the two whose product is represented by the lower right hand minor are coincident. Thus every cubic of the pencil is equianharmonic, with the exception of certain cuspidal cubics as degenerate cases.

(b) Second method for basis as in (a).

Using the same perspective R, we take as the nine base points of the pencil the intersections with $x_2 = 0$, $x_3 = 0$, $x_2 - x_3 = 0$ of a generic line and its correspondents under R, R^2:

$$\begin{aligned} a_1 x_1 + a_2 x_2 + a_3 x_3 &= 0, \\ \omega a_1 x_1 + a_2 x_2 + a_3 x_3 &= 0, \\ \omega^2 a_1 x_1 + a_2 x_2 + a_3 x_3 &= 0. \end{aligned}$$

The pencil is then

$$\lambda x_2 x_3 (x_2 - x_3) + (a_1 x_1 + a_2 x_2 + a_3 x_3)\ (\omega a_1 x_1 + a_2 x_2 + a_3 x_3) \times (\omega^2 a_1 x_1 + a_2 x_2 + a_3 x_3) = 0.$$

The Hessian is

$$\begin{vmatrix} 3a_1^3 x_1 & 0 & 0 \\ 0 & 3a_2^3 x_2 + x_3(3a_2^2 a_3 + \lambda) & (3a_2^2 a_3 + \lambda)x_2 + (3a_2 a_3^2 + \lambda)x_3 \\ 0 & (3a_2^2 a_3 + \lambda)x_2 + (3a_2 a_3^2 - \lambda)x_3 & (3a_2 a_3^2 - \lambda)x_2 + 3a_3^3 x_3 \end{vmatrix} = 0,$$

which is three non-concurrent lines unless the lower right hand minor is a perfect square, representing two coincident lines. The discriminant of this quadratic is of the second degree in λ. Thus two of the cubics of the pencil are cuspidal and the remainder are properly equianharmonic.

Kantor* makes a statement substantially as above if we recall his assertion that cuspidal cubics are to be included among equianharmonic cubics

* Kantor, S., Des transformations periodiques univoques, pp. 6, 9.

as degenerate cases. Chisini* states that Kantor is in error here and offers an example in contradiction. His example is a pencil constructed on a *nodal* cubic invariant under a cyclic perspective of order three and a triple of lines forming a cycle under this perspective, the presence of the nodal cubic purporting to contradict Kantor's statement. But a nodal cubic cannot be invariant under a cyclic perspective of the third order, since the node with its two tangents and the single flex line must be invariant elements. If the node is at the center and the flex line is the axis, the cubic degenerates since the three flex tangents meet at the node if they are individually invariant. If the three tangents form a cycle they must be concurrent at a point of the axis. In this case the node must lie in the axis, and the axis will be one nodal tangent. The other would go through the center. This is necessary since the nodal tangents must be individually invariant.

A cuspidal cubic may be invariant under a cyclic perspective of the third order, whose axis is the join of the cusp and the flex and whose center is the intersection of the cuspidal tangent and the flex tangent.

(c) Basis consisting of the three invariant points of a cyclic collineation of the third order as an in-circumscribed triangle, and three other points which form a cycle as simple base points.

Take the collineation $(x_1, \omega x_2, \omega^2 x_3)$ which has the invariant points (1, 0, 0), (0, 1, 0), (0, 0, 1). Cubics having this triangle as an in-circumscribed triangle have the form†

$$a_1 x_1^2 x_2 + a_2 x_2^2 x_3 + a_3 x_3^2 x_1 + a_4 x_1 x_2 x_3 = 0.$$

If we require this curve to pass through the cycle (ξ_1, ξ_2, ξ_3), $(\xi_1, \omega \xi_2, \omega^2 \xi_3)$, $(\xi_1, \omega^2 \xi_2, \omega \xi_3)$, we find that $a_4 = 0$, and the cubic becomes

$$a_1 x_1^2 x_2 + a_2 x_2^2 x_3 + a_3 x_3^2 x_1 = 0.$$

This cubic is clearly equianharmonic.

* Chisini, O., Sui gasci di cubiche a modulo costante, Rend. Circ. Mat. di Palermo, vol. 41 (1916), p. 91.

† Fixing the sense of the triangle.

URBANA, ILLINOIS.

ON FOUR MUTUALLY ORTHOGONAL CIRCLES.*

By Nathan Altshiller-Court.

1. **Introduction.** Four mutually orthogonal proper circles are considered in their relations to the triangles formed by their centers and the circumcircles of these triangles. To abbreviate the statement of some of the results obtained I shall refer to the triangle formed by the centers of three of the given circles and to the circumcircle of this triangle as the "central triangle" and the "central circle" related to the fourth given circle.

2. Lemma. *If three circles are mutually orthogonal, the centers of any two of them are a pair of conjugate points with respect to the third circle.*

The two given circles (B), (C) being orthogonal to each other, the circle (AB) described on their line of centers AB as diameter is coaxal with them. Now the third given circle (A) is, by assumption, orthogonal to (B) and (C), it is therefore also orthogonal to (BC), hence the ends B, C of its diameter BC are conjugate with respect to (A).†

3. If (A), (B), (C), (H) are four mutually orthogonal circles, the centers of any three of them are conjugate in pairs with respect to the fourth circle (2), hence: *If four circles are mutually orthogonal, each of them is the conjugate circle of the triangle formed by the centers of the remaining three given circles.*

4. As an immediate application of this proposition (3) we have: *Each of four mutually orthogonal circles is orthogonal to the orthoptic circle (or line) of any conic inscribed in its related central triangle.*‡

5. The circle (H) being the conjugate circle of the triangle ABC, the center H of (H) is the orthocenter of ABC, hence: *The centers of four mutually orthogonal circles form an orthocentric group of points.*§

6. The converse of the propositions (3) and (5) is true, namely: *The four conjugate circles of an orthocentric group of four triangles are mutually orthogonal.*

If $ABCH$ is an orthocentric group of points, the circles (BC), (AH) described on the segments BC, AH as diameters are orthogonal to each other. Now if (A) and (H) denote the conjugate circles of the triangles BCH, ABC, the points B, C are conjugate with respect to each of these

* Received November 1, 1927.

† Nathan Altshiller-Court, College Geometry, p. 153. Richmond, Va., 1925.

‡ Charles Michel, Compléments de Géométrie moderne, p. 32. Paris, 1926.

§ J. L. Coolidge, A treatise on the circle and the sphere, p. 141, theorem 11. Oxford, 1916.

two circles, and therefore (BC) is orthogonal to both (A) and (H). Thus the three circles (A), (H), (AH) have collinear centers and are orthogonal to the same circle (BC), hence these three circles are coaxal, and therefore (A) and (H) are orthogonal to each other.

7. The circumcircle (ABC) of the triangle ABC is the inverse, with respect to the conjugate circle (H) of ABC (3), of the nine point circle of this triangle.* Now the four triangles of the orthocentric group $ABCH$ have the same nine point circle, hence: *Given four mutually orthogonal circles, the inverse with respect to one of these circles of its related central circle is a fixed circle.*

8. The radius of the nine point circle (N) of ABC is equal to one half of the circumradius R of ABC. Now if u_h is the square of the radius of (H) and p_h is the power of H with respect to circumcircle (ABC) of this triangle, we have†

$$R/2 = R \cdot u_h/p_h \qquad \text{or} \qquad u_h = p_h/2.$$

Thus: *The square of the radius of each of four mutually orthogonal circles is equal to one half of the power of its center with respect to its related central circle.*

9. The foot D of the altitude AHD of the triangle ABC has for its polar with respect to the conjugate circle (H) of ABC the perpendicular to AHD at A. Thus the anticomplementary triangle $A'B'C'$ of ABC and the orthic triangle DEF of ABC are polar reciprocals with respect to (H), and the line EF is the polar of A' with respect to (H). Consequently the line $A'H$ is perpendicular to EF, and the foot X of $A'H$ on EF is the inverse of A' with respect to (H), so that the circle (XYZ) determined by X and its two analogous points Y, Z is the inverse of the circle $(A'B'C')$ with respect to (H). Now H is the incenter of DEF, the circle (XYZ) is therefore the incircle of DEF. On the other hand the circle $(A'B'C')$ has for its center the point H and for its radius the circumdiameter $2R$ of ABC. Thus we have, denoting by r the inradius of DEF,

$$u_h = -r \cdot 2R.$$

In a similar way we obtain the three analogous relations

$$u_a = r' \cdot 2R, \qquad u_b = r'' \cdot 2R, \qquad u_c = r''' \cdot 2R$$

where r', r'', r''' are the exradii of DEF, and u_a, u_b, u_c are the squares of the radii of the circle (A), (B), (C). Thus: *The squares of the radii*

* College Geometry, p. 217.

† Ibid., p. 208.

of four mutually orthogonal circles bear the same ratio to the respective radii of the four circles which are concentric with them and which touch the sides of the orthic triangle of the orthocentric group which consists of the centers of the four given circles.

This ratio is equal to the circumdiameter of any three of these last four points.

10. The preceding proposition (9) reduces the study of the relations between the squares of the radii of four mutually orthogonal circles to the study of the relations of the radii of the four circles touching the sides of a triangle. Every known relation of the latter, and they are many, may be paraphrased into a relation of the former. Here are two illustrations:* (a) *The sum of the reciprocals of the squares of the radii of four mutually orthogonal circles is zero.*†

(b) *The sum of the squares of the radii of four mutually orthogonal circles is equal to the square of the circumdiameter of the triangle formed by the centers of any three of these circles.*

11. The orthic axis $D'E'F'$ of ABC is the radical axis of (H) and (ABC), while BC is the radical axis of (H) and (A), hence $D' \equiv (D'E'F', BC)$ is a point on the radical axis of (ABC) and (A), the radical axis of these two circles is therefore the perpendicular from D' upon their line of centers OA, i. e., this radical axis coincides with the side $D'EF$ of the orthic triangle DEF of ABC. Hence: *Given four mutually orthogonal circles, consider the radical axes of any three of them with the circle passing through their centers. The twelve radical axes thus obtained coincide by fours in the three sides of the orthic triangle of the orthocentric group of points consisting of the centers of the given circles.*

12. The circle of similitude of the circles (B), (C) is coaxal with them, it is therefore orthogonal to both (A) and (H), hence its center lies on the radical axis EF (11) of these two circles. Thus: *Four given mutually orthogonal circles taken in pairs have six circles of similitude. The centers of these six circles lie on the sides of the orthic triangle of the orthocentric group of points consisting of the centers of the given circles.*

13. The three pairs of centers of similitude P, P'; Q, Q'; R, R' of the three pairs of circles $(B), (C)$; $(C), (A)$; $(A), (B)$ are the vertices of a complete quadrilateral having ABC for diagonal triangle. If we suppose that P', Q', R' are collinear, then the lines AP, BQ, CR meet in the trilinear pole U of the line $P'Q'R'$ with respect to the triangle ABC.

The circle of similitude (PP') of the circles (B), (C) has the segment PP' for diameter and is orthogonal to (H) (12), hence the points P, P' are

* College Geometry, articles 110 and 113.

† Coolidge, ibid., p. 139.

conjugate with respect to (H) and the line AP is the polar of P' with respect to (H). But the polars AP, BQ, CR of the points P', Q', R' with respect to (H) meet in the pole of $P'Q'R'$ with respect to (H), hence this pole coincides with U. Thus: *Given four mutually orthogonal circles, the trilinear pole of an axis of similitude of three of these circles with respect to the triangle determined by the centers of these three circles coincides with the pole of the same line with respect to the fourth given circle.*

14. It may be shown that this property (13) is characteristic of the axes of similitude of the given circles.

If M, m are trilinear pole and polar with respect to the triangle ABC, the traces I, I' of the lines m, AM on the line BC are separated harmonically by the points B, C. If M is at the same time the pole of m with respect to the conjugate circle (H) of ABC, the points I, I' are also conjugate with respect to (H), hence I, I' coincide with the points P, P', which proves the proposition.

THE UNIVERSITY OF OKLAHOMA.

A GENERAL THEOREM ON THE EXPRESSION OF A DETERMINANT IN TERMS OF ITS SUB-DETERMINANTS.*

By William W. Flexner.

It is the purpose of this paper to combine in a single, relatively simple proof several of the more fundamental theorems on determinants. The theorems included are:

1. Laplace's Expansion and its generalization.†

2. The theorem stating that the adjoint determinant is equal to the original determinant raised to the $(n-1)$st power.‡

3. The theorem giving the relation between a minor of an adjoint determinant and the corresponding cofactor of the original determinant,‡ and its generalization.§

In this paper repeated indices always indicate summation according to the Einstein summation convention. Use will frequently be made of the generalized Kronecker's delta,‖ $\delta^{i_1 i_2 \cdots i_k}_{a_1 a_2 \cdots a_k}$ where the indices may take on values from one to n. This symbol is zero if $a_i = a_j$ for any $i \neq j$ and if the upper indices are not some permutation of the lower. If the permutation is even, $\delta^{i_1 i_2 \cdots i_k}_{a_1 a_2 \cdots a_k} = +1$; if it is odd, $\delta^{i_1 i_2 \cdots i_k}_{a_1 a_2 \cdots a_k} = -1$. Now we define,

$$\varepsilon_{i_1 i_2 \cdots i_n} = \delta^{12\cdots n}_{i_1 i_2 \cdots i_n}, \qquad \varepsilon_{i_1 i_2 \cdots i_n} = \delta^{i_1 i_2 \cdots i_n}_{12\cdots n}.$$

Then an n-rowed determinant becomes,

$$\Delta = |g^j_i| = \varepsilon_{a_1 a_2 \cdots a_n} g^{a_1}_1 g^{a_2}_2 \cdots g^{a_n}_n$$

where the a's being repeated indices are summed each over its entire range of values, $1, 2, \cdots, n$, in accordance with the summation convention.‖ The subscripts on the elements of the determinant we shall agree to call column numbers and the superscripts row numbers. This distinction is made merely to facilitate presentation and is not necessary.

---

* Received November 1, 1927.

† Netto, "Erweiterung des Laplaceschen Determinantenzerlegungssatzes". Journ. f. Math. Bd. 114 (1895), pp. 145–352 and E. Pascal, "Die Determinanten" Leipzig, 1900, p. 99.

‡ Bocher, "Introduction to Higher Algebra", p. 31.

§ E. Pascal, loc. cit., p. 103, and Muir, Trans. Roy. Soc. Edinb., vol. 30 (1880–81), pp. 1–4.

‖ O. Veblen, "Invariants of Quadratic Differential Forms". Chap. 1, Cambridge, 1927.

A k-rowed minor containing the columns $(k+1)$ to $(k+p)$ and the rows $a_{k+1}, a_{k+2}, \cdots, a_{k+p}$ is defined as follows:

$$\delta^{a_{k+1}a_{k+2}\cdots a_{k+p}}_{b_{k+1}b_{k+2}\cdots b_{k+p}} g^{b_{k+1}}_{k+1} g^{b_{k+2}}_{k+2} \cdots g^{b_{k+p}}_{k+p}.$$

As the columns appearing in the minor are here assumed to be adjacent, this is not the most general definition; but it is general enough because any columns in a determinant may always be brought to a desired position with at most a change of sign. The rows appearing in the minor are left arbitrary. According to this definition, a minor may change sign with a change in the order of the rows, that is, a change in the order of $a_{k+1}, a_{k+2}, \cdots, a_{k+p}$. The classical definition assumes the rows and columns to be in the natural order. It therefore agrees with the new definition when $a_{k+1}, a_{k+2}, \cdots, a_{k+p}$ are arranged in order of ascending magnitudes.

The cofactor (or algebraic complement) of the p-rowed minor is defined as:

$$G^{k+1\cdots k+p}_{a_{k+1}\cdots a_{k+p}} = \varepsilon_{a_1\cdots a_n} g^{a_1}_1 \cdots g^{a_k}_k g^{a_{k+p+1}}_{k+p+1} \cdots g^{a_n}_n.$$

When $a_{k+1}, a_{k+2}, \cdots, a_{k+p}$ are in ascending order, this definition may be proved to agree with the classical one.

Having defined the quantities with which we are to work we can now come to the main theorem.

THEOREM: *Given the n-rowed determinant,*

$$\Delta = \varepsilon_{b_1 b_2 \cdots b_n} g^{b_1}_1 g^{b_2}_2 \cdots g^{b_n}_n$$

and an integer, m, $(m \leqq n)$, then

$$\delta^{a_1\cdots a_m}_{b_1\cdots b_m} G^{12\cdots m_1}_{a_1 a_2 \cdots a_{m_1}} G^{m_1+1\cdots m_1+m_2}_{a_{m_1+1}\cdots a_{m_1+m_2}} \cdots G^{m-m_k+1\cdots m}_{a_{m-m_k+1}\cdots a_m} = m_1!\, m_2! \cdots m_k!\, \Delta^{k-1} G^{12\cdots m}_{b_1 b_2 \cdots b_m}$$

where the G's are the cofactors in Δ of the minors with rows (lower indices) and columns (upper indices) as indicated and $m_1, m_2, \cdots, m_k$ are integers satisfying the single condition,

$$m_1 + m_2 + \cdots + m_k = m.$$

The meaning of the theorem may perhaps be made clearer by a statement in words. From an n-rowed determinant, Δ, take an m-rowed minor. Without loss of generality this minor may be assumed to contain the rows $1, 2, \cdots, m$ of Δ. Now divide the m-rowed minor determinant into matrices, each containing all m rows, but containing different numbers, $m_1, m_2, \cdots, m_k$ of columns. Then form the products of the determinants of these matrices as follows: each matrix and therefore each column, must be represented in

each product once and only once. Each row of the original determinant must be represented once and only once in each product. The totality of products must represent the totality of different combinations of the m_i-rowed determinants, $(i = 1, \cdots, k)$. Replace in each product each determinant by its cofactor in the original determinant Δ. The sum of these products, each with the proper sign, which is given by $\delta^{a_1 \cdots a_m}_{b_1 \cdots b_m}$, equals Δ^{k-1} times the cofactor in Δ of the m-rowed minor.

It now remains to account for the $m_1! m_2! \cdots m_k!$ which occur in the statement of the theorem. They are necessary because the summation of the a's introduces a repetition of each of the products of cofactors just discussed $m_1! m_2! \cdots m_k!$ times. This is because

$$G^{12\cdots m_1}_{a_1 a_2 \cdots a_{m_1}} = -G^{12\cdots m_1}_{a_2 a_1 \cdots a_{m_1}} = G^{123\cdots m_1}_{a_2 a_3 a_1 \cdots a_{m_1}} = \text{etc.}$$

which follows from the definition of cofactor.

The proof of the theorem in general case follows immediately from the proof in one special case by direct analogy. We shall make the proof in the case:

$$m_1 = 3, \quad m_2 = 2, \quad m_3 = m_4 = \cdots = m_k = 1.$$

In this case, then, $k = n - 3$.

The statement of the theorem becomes

$$\mathfrak{G} = \delta^{a_1 \cdots a_m}_{b_1 \cdots b_m} G^{123}_{a_1 a_2 a_3} G^{45}_{a_4 a_5} G^{6}_{a_6} \cdots G^{m}_{a_m} = 3!\, 2!\, \Delta^{n-4} G^{12\cdots m}_{b_1 b_2 \cdots b_m}.$$

Proof:

$$\begin{aligned} \mathfrak{G} = [&(\delta^{a_1 \cdots a_m}_{b_1 \cdots b_m})(\varepsilon_{a_1 a_2 a_3 i_4 \cdots i_n} g_4^{i_4} \cdots g_n^{i_n}) \\ &(\varepsilon_{j_1 j_2 j_3 a_4 a_5 j_6 \cdots j_n} g_1^{j_1} g_2^{j_2} g_3^{j_3} g_6^{j_6} \cdots g_n^{j_n}) \\ &(\varepsilon_{k_1 \cdots k_5 a_6 k_7 \cdots k_n} g_1^{k_1} \cdots g_5^{k_5} g_7^{k_7} \cdots g_n^{k_n}) \\ &\qquad \vdots \\ &(\varepsilon_{p_1 \cdots p_{m-1} a_m p_{m+1} \cdots p_n} g_1^{p_1} \cdots g_{m-1}^{p_{m-1}} g_{m+1}^{p_{m+1}} \cdots g_n^{p_n})]. \end{aligned} \tag{1}$$

The following simple relations exist between the generalized Kronecker symbols as may easily be verified.

$$\delta^{a_1 \cdots a_m}_{b_1 \cdots b_m} = \frac{1}{(n-m)!} \varepsilon^{a_1 \cdots a_m b_{m+1} \cdots b_n} \varepsilon_{b_1 \cdots b_m b_{m+1} \cdots b_n}, \tag{2}$$

$$\begin{aligned} &\delta^{a_1 \cdots a_k b_{k+1} \cdots b_{k+p} a_{k+p+1} \cdots a_m}_{c_1 \cdots \cdots \cdots c_m} \varepsilon_{b_1 \cdots b_n} \\ &\qquad = p!\, \delta^{a_1 \cdots a_k a_{k+p+1} \cdots a_m}_{c_1 \cdots c_k c_{k+p+1} \cdots c_m} \varepsilon_{b_1 \cdots b_p c_{k+1} \cdots c_{k+p} b_{k+p+1} \cdots b_n}. \end{aligned} \tag{3}$$

Therefore the first two brackets of (1) become

$$\left(3!\,\frac{\varepsilon_{b_1\cdots b_m b_{m+1}\cdots b_n}}{(n-m)!}\,\delta^{a_4\cdots a_m b_{m+1}\cdots b_n}_{i_4\cdots i_m i_{m+1}\cdots i_n}\,g_4^{i_4}\cdots g_n^{i_n}\right).$$

Now examining the other terms we see that each contains all but one or two of the column elements of Δ. For instance,

$$(\varepsilon_{j_1 j_2 j_3 a_4 a_5 j_6\cdots j_n} g_1^{j_1}\, g_2^{j_2}\, g_3^{j_3}\, g_6^{j_6}\cdots g_n^{j_n})$$

lacks $g_4^{a_4}$ and $g_5^{a_5}$. But if the members of $(g_4^{i_4}\cdots g_n^{i_n})$ be distributed among the parentheses, they will supply the necessary column elements to make a Δ in each case, with $(g_{m+1}^{i_{m+1}}\cdots g_n^{i_n})$ left over giving, when relation (3) is applied to $\delta^{a_4\cdots a_m b_{m+1}\cdots b_n}_{i_4\cdots i_m\, i_{m+1}\cdots i_n}$ and each of the ε's in turn:

$$\begin{aligned}\mathfrak{G} = [3!\,2!\,(&\varepsilon_{j_1 j_2 j_3 i_4 i_5 j_6\cdots j_n}\, g_1^{j_1}\cdots g_3^{j_3}\; g_4^{i_4}\; g_5^{i_5}\; g_6^{j_6}\cdots g_n^{j_n})\\ &(\varepsilon_{k_1\cdots k_5 i_6 k_7\cdots k_n}\, g_1^{k_1}\cdots g_5^{k_5}\; g_6^{i_6}\; g_7^{k_7}\cdots g_n^{k_n})\\ &\qquad\vdots\\ &(\varepsilon_{p_1\cdots p_{m-1} i_m p_{m+1}\cdots p_n}\, g_1^{p_1}\cdots g_{m-1}^{p_{m-1}}\, g_m^{i_m}\, g_{m+1}^{p_{m+1}}\cdots g_n^{p_n})\\ &(\varepsilon_{b_1\cdots b_m i_{m+1}\cdots i_n}\, g_{m+1}^{i_{m+1}}\cdots g_n^{i_n})].\end{aligned}$$

The last expression reduces to:

$$\mathfrak{G} = 3!\,2!\,\Delta^{n-4}\, G^{1\,2\cdots\cdots m}_{b_1 b_2\cdots b_m}.$$

The general case where arbitrary $m_1, m_2, \cdots, m_k$ are used, goes through in exactly the same way.

The theorems mentioned at the beginning of the paper come out as follows:

1. Laplace's Expansion: $m = n$, $k = 2$; and its generalization: $m = n$, $k \neq n$.

2. Adjoint determinant $= \Delta^{n-1}$: $m = n$, $k = n$.

3. Theorem relating minor of adjoint and cofactor of Δ: $m \neq n$, $k = n$; and its generalization: $m \neq n$, $k \neq n$.

QUADRATIC FIELDS IN WHICH CYCLOTOMIC POLYNOMIALS ARE REDUCIBLE.*

By Louis Weisner.

1. **Introduction.** Gauss† was the first to prove that the cyclotomic polynomial $F_m(x)$ of degree $\varphi(m)$, whose zeros are the primitive mth roots of unity, is reducible in $R\left(\sqrt{(-1)^{(m-1)/2}\, m}\right)$, m being an odd prime. Dirichlet‡ extended this theorem to the case in which m is the product of two distinct odd primes, treating as an example $m = 33$. Cauchy§ stated without proof that the theorem could be extended to the case where $m = 2^a p_1 p_2 \cdots p_r$, the p's being distinct odd primes and $a = 0$, 2 or 3. Genocchi|| justified Cauchy's remarks, but the problem of factoring F_m in quadratic fields was not settled thereby, as there are, in general, other quadratic fields besides that mentioned, in which F_m is reducible. In the present paper all such fields are determined, the statement of results being deferred to § 7.

Problems of this sort are conveniently treated by means of Galois' theory of equations. If $f(x) = 0$ is an irreducible equation with coefficients in R and G its group relative to R, G is transitive. If, on adjoining $\sqrt{k}$ to R, where k is in R, the group reduces to an intransitive subgroup of G, $f(x)$ is reducible in R, and not otherwise. Conversely, to each intransitive subgroup G' of G of index 2, there exists a quadratic field (quadratic relative to R) in which $f(x)$ is reducible. This quadratic field may be found by computing the numerical value of some function belonging to G'. As no two distinct subgroups of index 2 can lead to the same quadratic field, the number of quadratic fields in which $f(x)$ is reducible is equal to the number of intransitive subgroups of G of index 2.

For a particular value of m the preceding method may be employed to determine all quadratic fields in which F_m is reducible.¶ For a general m it hardly seems practicable to compute the numerical values of the functions belonging to the intransitive subgroups of index 2 of the group G of $F_m = 0$

* Received November 8, 1927. Presented to the American Mathematical Society October 29, 1927.

† C. F. Gauss, *Disquisitiones Arithmeticae* (1801), p. 637; *Werke*, vol. 1, p. 443.

‡ G. L. Dirichlet, *Sur la manière de resoudre l'equation* $t^2 - pu^2 = 1$ *au moyen des fonctions circulaires*, Jour. für Math., vol. 17 (1837), p. 270; *Werke*, vol. 1, p. 350.

§ A. L. Cauchy, *Œuvres*, ser. 1, vol. 5, p. 84.

|| A. Genocchi, *Intorni ad un teorema di Cauchy*, Annali di Mat., (2), vol. 2 (1868–9), p. 216.

¶ Cf. H. Weber, *Algebra*, vol. 2 (1899), p. 73–6.

relative to $R(1)$. We avoid these calculations by proving two lemmas from which we infer, with the aid of the theorem of Gauss referred to, that F_m is reducible in certain quadratic fields. On computing the number of these fields we find that it equals the number of intransitive subgroups of G of index 2, and hence conclude that we have found all the quadratic fields in which F_m is reducible.

2. **Group of the cyclotomic equation $F_m = 0$.** Let ε be a primitive mth root of unity and $c_1 = 1, c_2, \cdots, c_{\varphi(m)}$ the positive integers less than m and prime to m. The cyclotomic polynomial F_m is the expansion of

$$\prod_{i=1}^{i=\varphi(m)} (x - \varepsilon^{c_i}).$$

$F_m = 0$ is an abelian equation and is irreducible in $R(1)$. Its group G relative to $R(1)$ is therefore a regular abelian group of order $\varphi(m)$. If s_i is the substitution of G which replaces ε by ε^{c_i}, s must replace ε^{c_j} by $\varepsilon^{c_i c_j}$ as every substitution of G must leave

$$\varepsilon^{c_j} - (\varepsilon)^{c_j}$$

unaltered in value. The $\varphi(m)$ substitutions of G therefore consist of

$$s_i \equiv \begin{pmatrix} \varepsilon, & \varepsilon^{c_2}, & \cdots, & \varepsilon^{c_{\varphi(m)}} \\ \varepsilon^{c_i}, & \varepsilon^{c_i c_2}, & \cdots, & \varepsilon^{c_i c_{\varphi(m)}} \end{pmatrix}, \qquad (i = 1, 2, \cdots, \varphi(m)).$$

*G is simply isomorphic with the group formed by the totitives $1, c_2, \cdots, c_{\varphi(m)}$ of m with respect to multiplication modulo m.** For if we make s_i correspond to c_i, $s_i s_j$ corresponds to $c_i c_j$, since $s_i s_j$ replaces ε by $\varepsilon^{c_i c_j}$.

3. LEMMA 1. *If F_m is reducible in the distinct quadratic fields $R(\sqrt{k_1})$, $R(\sqrt{k_2})$, it is also reducible in $R(\sqrt{k_1 k_2})$.* (The latter field is necessarily distinct from the other two.)

The numbers $\sqrt{k_1}$, $\sqrt{k_2}$ equal rational functions of ε with coefficients in $R(1)$; let these functions be f_1, f_2 respectively, and suppose them to belong to the subgroups H_1, H_2 respectively, each of which is necessarily of index 2 in G. If I is the cross-cut of H_1 and H_2, G/I is a non-cyclic group of order 4. Hence there are only three subgroups of G besides G that contain I. Two of these we know to be H_1 and H_2; let H_3 be the third.

The function $f_1 f_2$, being two-valued and invariant under I, belongs to one of these three groups. Now $f_1 f_2$ is not invariant under H_1, for in

* *Loc. cit.*, p. 74.

that case f_2 would be invariant under H_1. Similarly $f_1 f_2$ is not invariant under H_2. It follows that $f_1 f_2$ belongs to H_3. Hence H_3 is the group of $F_m = 0$ relative to $R(\sqrt{k_1 k_2})$. As H_3, as well as every subgroup of G, is intransitive, F_m is reducible in $R(\sqrt{k_1 k_2})$.

4. LEMMA 2. *If m' is a multiple of m, $F_{m'}$ is reducible in every quadratic field in which F_m is reducible.*

Let F_m be reducible in $R(\sqrt{k})$ and $k = f(\epsilon)$, the coefficients of f being in $R(1)$. If H is the subgroup of G to which f belongs, there is a composition series of G whose first two terms are G, H. If we proceed to solve $F_m = 0$ by means of this composition series in the usual way, the first step consists of the adjunction of $\sqrt{k}$ to the rational domain. The final result exhibits ϵ in terms of radicals, the first root extraction being $\sqrt{k}$.

Conversely, if ϵ is expressed in terms of radicals in which the first root extraction is $\sqrt{k}$, this particular expression for ϵ is obtainable by solving $F_m = 0$ by means of some composition series of G whose first two terms are G, H; hence F_m is reducible in $R(\sqrt{k})$. Now $\sqrt[m/m']{\epsilon}$ is an expression in terms of radicals of a primitive m'th root of unity whose first root extraction is also $\sqrt{k}$. Hence $F_{m'}$ is reducible in $R(\sqrt{k})$.

5. **Certain fields in which F_m is reducible.** Let $m = 2^{a_0} p_1^{a_1} \cdots p_r^{a_r}$, where $p_1, \cdots, p_r$ are distinct odd primes. From lemma 2 and the theorem of Gauss stated in § 1, we infer that F_m is reducible in $R(\sqrt{(-1)^{(p_i-1)/2} p_i})$, $(i = 1, \cdots, r)$. From lemma 1 we infer that F_m is reducible in $R(\sqrt{k})$, where k is the product of any $r' \leqq r$ of the numbers $(-1)^{(p_i-1)/2} p_i$. The number of quadratic fields thus found in which F is reducible is

$$\binom{r}{1} + \binom{r}{2} + \cdots + \binom{r}{r} = 2^r - 1.$$

If $a_0 \geqq 2$, $F_{a_0} = x^{2^{a_0-1}} + 1$ is reducible in $R(\sqrt{-1})$. Hence F_m is reducible in $R(\sqrt{-1})$ and in $R(\sqrt{\pm k})$, where k is the product of any $r' \leqq r$ of the p_i's. The number of quadratic fields in which F_m is reducible is increased to

$$2(2^r - 1) + 1 = 2^{r+1} - 1.$$

Finally, if $a_0 \geqq 3$,

$$F_{a_0} \equiv x^{2^{a_0-1}} + 1 = (x^{2^{a_0-2}} + \sqrt{2}\, x^{2^{a_0-3}} + 1)(x^{2^{a_0-2}} - \sqrt{2}\, x^{2^{a_0-3}} + 1)$$

is reducible in $R(\sqrt{2})$, and hence also in $R(\sqrt{\pm 2k})$, as well as in $R(\sqrt{-1})$, $R(\sqrt{-2})$ and $R(\sqrt{\pm k})$, where k again is the product of

any $r' \leqq r$ of the p_i's. The number of quadratic fields in which F_m is reducible is increased to

$$1+2\left[\binom{r+1}{1}+\binom{r+1}{2}+\cdots+\binom{r+1}{r+1}\right]=2^{r+2}-1.$$

Summarizing, we have found 2^r-1 quadratic fields in which F_m is reducible if m is odd or twice an odd number; $2^{r+1}-1$ if m is 4 times an odd number; and $2^{r+2}-1$ if m is a multiple of 8.

6. **Intransitive subgroups of G of index 2.** The number of intransitive subgroups of G of index 2 is clearly equal to the number of subgroups of index 2 of the group T of the totitives of m. To find this number we employ the basis $b_1, \cdots, b_m$ of T given in Weber's *Algebra*, vol. 2 (1899), p. 66–8. The number n of elements of the basis is r, $r+1$, or $r+2$ according as m is odd or twice an odd number, 4 times an odd number, or a multiple of 8, r being the number of distinct odd prime factors of m. The order of every element of the basis is even.

Every subgroup of T of index 2 contains the subgroup I consisting of $1, c_2^2, \cdots, c_{\varphi(m)}^2$, which is clearly generated by $b_1^2, \cdots, b_n^2$. If I contained b_1, we should have

$$b_1 \equiv (b_1^{e_1}\, b_2^{e_2} \cdots b_n^{e_n})^2 \mod m,$$

where $e_1, e_2, \cdots, e_n$ are positive integers or zero. Hence

$$b_1^{2e_1-1} \equiv 1, \quad b_2^{2e_2} \equiv 1, \cdots, b_n^{e_n} \equiv 1 \mod m.$$

As the first of these congruences asserts that the order of b_1 is odd, I does not contain b_1, and similarly I does not contain $b_2, \cdots, b_n$.

Hence T/I is generated by its elements $\beta_1, \beta_2, \cdots, \beta_n$ corresponding to $b_1, b_2, \cdots, b_n$ respectively. Furthermore $\beta_1, \beta_2, \cdots, \beta_n$ form a set of *independent* generators of T/I; for if, for example,

$$\beta_1 = \beta_2^{h_2} \cdots \beta_n^{h_n}$$

then

$$b_1 \equiv b_2^{h_2} \cdots b_n^{h_n}\, (b_1^{e_1}\, b_2^{e_2} \cdots b_n^{e_n})^2 \mod m,$$

whence

$$b_1^{2e_1-1} \equiv 1, \quad b_2^{h_2+2e_2} \equiv 1, \cdots, b_n^{h_n+2e_n} \equiv 1 \mod m.$$

Here again the first congruence implies that the order of b_1 is odd, contrary to fact. Hence $\beta_1, \beta_2, \cdots, \beta_n$ form a set of independent generators of T/I.

As the square of every element of T is in I, T/I is an abelian group of type $(1, 1, \cdots, 1)$ and order 2^n. The number of subgroups of index 2

of this group is known to be 2^n-1,* which is also the number of subgroups of T of index 2. Thus T and G contain exactly 2^r-1, $2^{r+1}-1$, or $2^{r+2}-1$ subgroups of index 2, according as m is odd or twice an odd number, 4 times an odd number, or a multiple of 8.

7. Conclusion. Comparing the results § 5 with those of § 6, we conclude that the quadratic fields of § 5 are the only ones in which F_m is reducible. Hence the

THEOREM. *F_m is reducible in $R(\sqrt{d})$ and in no other quadratic fields, d extending over all positive and negative divisors of m (except $+1$) which are not divisible by the square of a prime, subject to the following restrictions*: (1). *If m is 4 times an odd number, d must be odd.* (2). *If m is odd or twice an odd number, d must be odd and $\equiv 1 \bmod 4$.*

* Cf. W. Burnside, *Theory of Groups* (1911), p. 110.

HUNTER COLLEGE OF THE CITY OF NEW YORK.

ON IRREDUCIBLE CUTS OF THE PLANE BETWEEN TWO POINTS.†

By Wallace Alvin Wilson.

1. A set F is said to be an irreducible cut (*coupure irreductible*) between two points of its complement, if no proper closed sub-set of F separates the points. The set F is called a completely irreducible cut (*coupure irreductible complète*) of the plane if it separates two points of its complement but no proper closed part separates any two points of its complement.

The following theorems regarding irreducible cuts have been proved by C. Kuratowski:‡

(a) For a closed set F to be an irreducible cut between two points m and n it is necessary and sufficient that, if R_m and R_n are the components of the complement of F containing m and n, respectively, then F is the frontier of both R_m and R_n.

(b) A bounded completely irreducible cut is either an indecomposable continuum or it is the union of two continua each of which is irreducible between any pair of points belonging to different components of their divisor.

From Theorem (a) it follows that a completely irreducible cut is the frontier of *every* component of its complement and conversely. It is easily seen that Theorem (b) becomes invalid if the word "completely" is omitted, or, what amounts to the same thing, if the set is the frontier of two or more, but not all of the components of its complement. It is the purpose of the present article to derive a theorem analogous to (b) for the case that the cut is irreducible, but not necessarily completely so. The principal theorems are found in §§ 5 and 9.

2. Notation. If F is an irreducible cut between the points m and n of its complement, we shall say that "F is an irreducible $S(m, n)$".

If A is a proper part of B, the set of inner points of A relative to B will be denoted by A^*.

The whole plane is denoted by Z.

If x is a point, the symbol $V_\delta(x)$ denotes the interior of a circle of center x and radius δ.

If F is an irreducible $S(m, n)$, the components of $Z - F$ containing m and n will be denoted by R_m and R_n, respectively.

A component of $Z - F$ whose frontier is F will be called a *principal* component; others will be called *secondary*.

† Received September 16, 1927.

‡ (A) C. Kuratowski, Sur les coupures du plan, Fundamenta Math., vol. 6, pp. 130–145. See also R. L. Moore, Concerning the common boundary of two domains, ibid., pp. 203–213.

3. THEOREM. *Let F be a bounded decomposable irreducible $S(m, n)$. Then there is an irreducible decomposition† of F into proper sub-continua H and K such that*: (a) $H^* = F - K$ *and* $K^* = F - H$ *are connected*; (b) $\overline{H^*} = H$, $\overline{K^*} = K$; (c) H *and* K *are irreducible about* $H \cdot K$.

Proof. Since F is decomposable, it contains a proper sub-continuum M which is not a continuum of condensation. Then $F - M$ is connected‡ and so $\overline{F - M}$ is a proper sub-continuum of F and not a continuum of condensation. Set $H = \overline{F - M}$ and $K = \overline{F - H}$. Then $F = H + K$ is a decomposition of F into proper sub-continua. It is irreducible because $K = \overline{F - H}$ and $H = \overline{F - K}$ by a theorem of Kuratowski.‡ Thus (a) and (b) hold.

Suppose now that H' is a proper sub-continuum of H containing $H \cdot K$. Then $H' + K$ is not an $S(m, n)$. Also H is not an $S(m, n)$. But $H(H' + K) = H'$. Then by a theorem essentially due to Janiszewski§ we have the contradiction that $H + K$ is not an $S(m, n)$. Hence H is irreducible about $H \cdot K$ and in like manner so is K.

4. LEMMA. *Let F be a bounded irreducible $S(m, n)$. Let A and B be proper sub-continua of F which are not continua of condensation and let $A \cdot B = 0$. Then F is the union of two proper sub-continua which are irreducible about $A + B$ and have no other points in common.*

Proof. There exists by a theorem of Miss Mullikin‖ a sub-continuum P of F which is irreducible between A and B. Then $H = A + P + B$ is irreducible about $A + B$ and obviously $H \neq F$ by reference (B).

Set $Q = \overline{F - H}$. It is easily seen that $P \cdot Q = 0$ and that neither $Q \cdot A$ nor $Q \cdot B$ is void. Hence Q contains a sub-continuum Q' irreducible between A and B. If Q' does not contain $F - H$, we have a contradiction. For neither $Q' + H$ nor $Q + A + B$ is an $S(m, n)$, but $(Q + A + B)(Q' + H) = A + B + Q'$, a continuum, and $(Q + A + B) + (Q' + H) = F$. Thus $Q' \supset F - H$ and hence $Q' = Q$.

Set $K = A + Q + B$. Then K is irreducible about $A + B$ and $H \cdot K = A + B$. Hence H and K satisfy the requirements of the theorem.

5. THEOREM. *Let F be a bounded irreducible $S(m, n)$ containing two sub-continua A and B which have no common points and are not continua of condensation. Then F is the union of four continua, L, M, P, Q, having these properties*: (a) *each contains points not in any of the other three*;

† A decomposition of F into sub-continua is called irreducible if F is not the union of one of them and a proper sub-continuum of the other.

‡ See reference (A), p. 136.

§ (B) Z. Janiszewski, Sur les coupures du plan faites par des continus, Prace matem.-fizyczne, vol. 26, Theorem A: If P and Q are bounded closed sets, if $P \cdot Q$ is connected, and neither P nor Q is an $S(m, n)$, then $P + Q$ is not an $S(m, n)$.

‖ (C) Anna M. Mullikin, Certain theorems relating to plane connected sets, Trans. Amer. Math. Soc., vol. 24, p. 145.

(*b*) $L \cdot M = P \cdot Q = 0$; (*c*) *$L$ and M are irreducible between P and Q*; (*d*) *P and Q are irreducible between L and M.*

Proof. Let P and Q be as defined in § 4 and let L and M be sub-continua of A and B, respectively, irreducible between P and Q. It was shown in § 4 that P and Q are irreducible between A and B and hence, a fortiori, between L and M. This gives (b), (c), and (d). It is easily seen that property (a) is also true.

It remains to show that $F = L+M+P+Q$. By § 4, $F = A+B+P+Q$. If L did not contain all points of A not in $B+P+Q$, we would have a contradiction. For $L+B+P+Q$ would not be an $S(m, n)$ and $A+P+Q$ is not an $S(m, n)$, but $(L+B+P+Q)(A+P+Q) = L+P+Q$, a continuum. Hence $F = L+B+P+Q$ and a repetition of this reasoning gives $F = L+M+P+Q$.

COROLLARY 1. *Let F be a bounded irreducible $S(m, n)$ containing sub-continua A and B which have no common points and are not continua of condensation. Then there is an irreducible decomposition of F into continua H and K, such that $H \cdot K = \alpha + \beta$, where α and β are closed sets, $\alpha \cdot \beta = 0$, and H and K are irreducible between α and β.*

Proof. Using the notation of the theorem, set $H = L+Q$, $K = M+P$, $\alpha = L \cdot P$, $\beta = M \cdot Q$. Then $\alpha \cdot \beta = 0$, since $L \cdot M = 0$. If a sub-continuum H' of H joins α to β, it joins P to Q, and so contains L or M. As M is not a part of H, H' contains L. Likewise, $H' \supset Q$. Hence $H' = H$ and H is irreducible between α and β. So also is K.

If the decomposition were not irreducible, we would have a sub-continuum H' of H such that $F = H' + K$. If $H' \cdot M$ or $H' \cdot P$ were void, there would be an immediate contradiction by reference (B). Hence, $H' \cdot \alpha \neq 0$ and $H' \cdot \beta \neq 0$, which is impossible unless $H' = H$, since H is irreducible between α and β.

COROLLARY 2. *Let F be a bounded irreducible $S(m, n)$ and $F = A+B+C+D$ be an irreducible decomposition of F into four continua such that $A \cdot C = B \cdot D = 0$. Then A and C are irreducible between B and D, and B and D between A and C.*

Remarks. It is easily seen that the decomposition $F = L+M+P+Q$ in the theorem is irreducible. It should be noted that Corollary 1 is an analogue of Kuratowski's theorem (See § 1) for a bounded irreducible $S(m, n)$, which contains two sub-continua that have no common points and are not continua of condensation. The remaining sections of this article enable us to omit this restriction. In the restricted case there are exactly two principal components of the complement of F. This follows from a generalization of Rosenthal's theorem.†

† (D), W. A. Wilson, On the separation of the plane by irreducible continua, Bull. Amer. Math. Soc., vol. 33, p. 739.

6. It was shown in § 3 that, if F is a decomposable bounded irreducible $S(m, n)$, there is an irreducible decomposition of F into continua H and K with the properties there enumerated. The connected sets $H^* = F - K$ and $K^* = F - H$ are separated sets (i. e., $\bar{H}^* \cdot K^* + H^* \cdot \bar{K}^* = 0$) and so, by a theorem of Knaster and Kuratowski† there is, for any points a of H^* and b of K^*, a bounded continuum E which is an $S(a, b)$ and does not meet $H^* + K^*$.

By a theorem of Mazurkiewicz‡ there is a sub-continuum P of E which is an irreducible $S(a, b)$. Since H^* and K^* are connected, P is an irreducible $S(H^*, K^*)$, i. e., P cuts the plane between every pair of points of which one lies in H^* and the other in K^*. Since $\bar{H}^* = H$ and $\bar{K}^* = K$, $H \cdot K \subset P$. It should also be noted that P contains points in every principal component of $Z - F$.

LEMMA. *Let F be a bounded irreducible $S(m, n)$ and $F = H + K$ an irreducible partition of F into proper sub-continua. Let P be any bounded irreducible cut between $H^* = F - K$ and $K^* = F - H$. Then, if R is a principal component of $Z - F$, there is a continuum Q which is an irreducible $S(H^*, K^*)$ and is composed of an arc A lying in R and a sub-continuum C of P which is irreducible between the ends of A and meets A in no other point.*

Proof. Let m be a point of P in R. Let G' and G'' be the components of $Z - P$ containing K^* and H^*, respectively. About m draw a circumference J within R and let I be its interior. Since P is the frontier of G' and G'', it is easy to see that there is a simple arc xy in G' and a uv in G'' joining points x and u of J to points y and v, respectively, of K^* and H^* and meeting J in no other points. Draw the radii mx and mu and set $vy = uv + mu + mx + xy$. Let uax and ubx be the arcs into which u and x divide J.

Let $P_1 = P - P \cdot I$. Then $P_1 + J$ is an $S(H^*, K^*)$. For $\bar{I}$ is not an $S(H^*, K^*)$ and $(P_1 + J)\bar{I} = J$, a continuum, but $P_1 + \bar{I}$ is an $S(H^*, K^*)$. (See reference (B).)

Moreover, neither $P_1 \cdot (uax)$ nor $P_1 \cdot (ubx)$ is void. For otherwise an application of Janiszewski's theorem would give a contradiction. Let $u'x'$ and $u''x''$ be arcs of uax and ubx, respectively, irreducible about $P_1 \cdot (uax)$ and $P_1 \cdot (ubx)$. Then P_1 contains a sub-continuum C irreducible between $u'x'$ and $u''x''$ or else $P_1 = p + q$, where p and q are closed and $p \cdot q = p \cdot (u''x'')$

† (E) B. Knaster und C. Kuratowski, Sur les ensembles connexes, Fundamenta Math. vol. 2, p. 233. "If the sets A and B are separated, at least one being bounded, and a and b are any points belonging to A and B respectively, there exists a continuum not meeting $A + B$ and cutting the plane between a and b."

‡ (F) S. Mazurkiewicz, Sur un ensemble G_δ, etc., Fundamenta Math., vol. 1, p. 63.

$= q \cdot (u'x') = 0$ by a well-known theorem. The latter case gives a contradiction.

Then, reasoning as above, we see that $C+J$ is an $S(H^*, K^*)$. For $(C+J)(P_1+u'x'+u''x'') = C+u'x'+u''x''$. Let A be an arc of J irreducible between $C \cdot (u'x')$ and $C \cdot (u''x'')$. A repetition of the reasoning just used shows that $Q = A+C$ is an $S(H^*, K^*)$. Obviously C is irreducible between the ends of A and meets A in no other points. By using the facts that A and C are irreducible continua and Janiszewski's theorem, it is easily seen that Q is an irreducible $S(H^*, K^*)$.

7. The essential effect of the lemma of the previous section is to replace a portion of any irreducible $S(H^*, K^*)$ by an arc lying in one of the principal components of $Z-F$. This can be repeated for each principal component.

Let us denote the sets R, Q, A, and C of § 6 by R_1, Q_1, A_1, and C_1. Let R_2 be another principal component of $Z-F$. Then by the lemma there is an irreducible cut $Q_2 = A_2+C_2$ between H^* and K^*, where A_2 is an arc in R_2 and C_2 is a sub-continuum of Q_1 irreducible between the ends of A_2 and meeting A_2 in no other point. Now $C_2 \supset A_1$, since otherwise H^* and K^* could be joined by an arc in R_1. As A_1 is not a continuum of condensation and $A_1 \cdot A_2 = 0$, $\overline{C_2 - A_1} = B_1+B_2$, where $B_1 \cdot B_2 = 0$ and B_1 and B_2 are continua, each of which is irreducible between an end of A_1 and an end of A_2. Also $H \cdot K = F \cdot Q_2 = F \cdot B_1 + F \cdot B_2$, and neither $F \cdot B_1$ nor $F \cdot B_2$ is void, as F separates A_1 from A_2.

If R_3 is another principal component of $Z-F$, it contains points of either B_1 or B_2, say of B_2. Then as above we see that there is an irreducible cut $Q_3 = A_3+C_3$, which is obtained from Q_2 by replacing B_2 by an arc in R_3 and two sub-continua B_2' and B_3' of B_2, such that $B_2' \cdot B_3' = 0$, B_2' is irreducible between an end of A_2 and one of A_3, and B_3' is irreducible between an end of A_3 and one of A_1. As before, $H \cdot K = F \cdot Q_3 = F \cdot B_1 + F \cdot B_2' + F \cdot B_3'$, and none of these is void.

If this process is continued, we obtain a sequence of irreducible cuts $\{Q_k\}$ between H^* and K^*, which will be finite or infinite according as the number of principal components of $Z-F$ is finite or infinite. Each Q_k is composed of k arcs $\{A_i\}$ and k continua $\{B_i\}$, where $A_i \cdot A_j = B_i \cdot B_j = 0$, $A_i \subset R_i$ for each i, $B_i \cdot A_j = 0$ unless $j = i$ or $j = i+1 \bmod k$, each B_i is irreducible between an end of A_i and one of A_{i+1} and meets these arcs in no other points, $H \cdot K = F \cdot Q_k = \sum_1^k F \cdot B_i$ and $\sum_1^{k-1} A_i + \sum_1^k B_i \subset Q_{k-1}$. To save words such a sequence of irreducible cuts between H^* and K^* will be called *normal* and any one of the cuts Q_k will be called a *normal* $S(H^*, K^*)$.

8. If R is a simply connected region whose frontier F is bounded and A and C are subsets of F without common points, we say that the set H *separates* A *from* C *in* R provided that $H \subset \bar{R}$ and every continuum in $\bar{R}$ which does not meet $F-(A+C)$ but joins A and C meets H. We shall say for brevity that H is an $S(A, C)$ in R. If H is an $S(A, C)$ in R, but no closed part of H is one, then H is an irreducible $S(A, C)$ in R. (These definitions are modifications of certain ones used by S. Mazurkiewicz. See reference (F), p. 62.)

If A and C have inner points relative to F, it is easily seen that a necessary and sufficient condition for a bounded closed sub-set H of $\bar{R}$ to separate A^* from C^* in R is that every simple arc in $\bar{R}$ joining A^* to C^* but meeting F in only the end-points have a point in common with H.

LEMMA. *Let F be a bounded irreducible $S(m, n)$ and $F = H+K$ be an irreducible decomposition of F into proper sub-continua. Let P be a normal $S(H^*, K^*)$ and $P = A+B+C+D$, where A and C are arcs in R_m and R_n, respectively, and $F \cdot P = F \cdot B + F \cdot D$. Then H and K are irreducible between B and D.*

Proof. Let G' and G'' be the components of $Z-P$ containing H^* and K^*, respectively, and let H' be a proper closed part of H. Then there is a point x of H^* not in H'. For some $\delta > 0$ there is a $V_\delta(x)$ in G' such that $H' \cdot V_\delta(x) = 0$. Let m lie in $R_m \cdot V_\delta(x)$ and n in $R_n \cdot V_\delta(x)$. Since n lies in G' and C^* is an arc, there is a simple arc nz, such that z lies in C^*, $nz - z \subset G'$, and $(nz)(B+A+D+K) = 0$. Since $n+z \subset R_n$, $H+K$ is not an $S(n, z)$. As $(H+K)(B+A+D+K) = K$, $H+K+B+A+D$ is not an $S(n, C^*)$. Hence there exists an arc nz not cutting any of these sets such that $C^* \cdot (nz) = z$ and $nz - z \subset G'$.

Since P is an irreducible $S(H^*, K^*)$, $B+C+D$ is not an $S(m, K^*)$ and m can be joined to K^* by a simple arc not cutting $B+C+D$. Starting from m, let u be the first point where this arc meets K^* and let w lie on mu and be in some $V_\eta(u)$ so small that $P \cdot V_\eta(u) = 0$. As w is in G'', $(mw) \cdot A^* \neq 0$. Let mw meet A for the last time at m'. Then $m'w \subset \bar{G}''$ and $m'w \subset R_m$ since $A \subset R_m$. That is, w lies in R_m. Thus $B+C+D+K$ is not an $S(m, w)$; nor is $H+K$, since $m+w \subset R_m$. Hence $H+K+B+C+D$ is not an $S(m, w)$, and there is an arc mw not cutting any of these sets. Starting from m let y be the first point where it meets A^*. Then $my \subset \bar{G}'$ and $(my) \cdot H' = 0$.

Let mn be an arc in $V_\delta(x)$. Then $(mn) \cdot H' = 0$ and $yz = my+mn+nz$ is an arc in $\bar{G}'$ not meeting H' and meeting P only in the points y and z in A^* and C^*, respectively. Thus no closed part of H is an $S(A^*, C^*)$ in $\bar{G}'$.

Let M be a sub-continuum of H irreducible between B and D. Any simple arc xy in $\bar{G}'$ which meets P in only two points, x in A^* and y

in C^*, divides G' into two connected regions such that the frontier of each contains one of the sets B and D, but not the other. As $M-M\cdot(B+D)$ is connected and contains points in both these regions and $M\cdot(A+C)=0$, M meets xy. Hence M is an $S(A^*, C^*)$ in $\overline{G}'$. This with the conclusion of the previous paragraph shows that H is irreducible between B and D. A similar proof is valid for K.

9. THEOREM. *Let F be a bounded decomposable irreducible $S(m, n)$ and let $F=H+K$ be an irreducible decomposition of F into proper sub-continua. If there is a finite number n (or an infinite number) of principal components of the complement of F, $H\cdot K$ is the sum of at least n (or an infinite number) of closed sets without common points such that H and K are both irreducible between each pair.*

Proof. Let $\{Q_k\}$ be the normal sequence of irreducible cuts between H^* and K^* obtained in § 7. We first take the case that $Z-F$ has a finite number n of principal components $\{R_i\}$. Then there is a $Q_k=Q_n$.

Using the notation and results of § 7, we see that for any $k\leqq n$, $H\cdot K$ is the sum of k closed sets $\{\alpha_i\}$, where each $\alpha_i=F\cdot B_i$ and $\alpha_i\cdot\alpha_j=0$. Consider any one of these, as α_1. Now $Q_k=A_1+B_1+A_2+M$, where $M\cdot B_1=0$, A_1 and A_2 are arcs in R_1 and R_2, respectively, $M=\sum_2^k B_i+\sum_3^k A_i$, and M and B_1 are irreducible between the ends of A_1 and A_2. Set $\mu=\sum_2^k\alpha_i$; then $\mu\subset M$.

By § 8 H and K are irreducible between α_1 and μ. A fortiori, H and K are irreducible between α_1 and each α_i, $i\neq 1$. Thus H and K are irreducible between each pair of sets α_i, α_j, $i\neq j$. Taking $k=n$, this proves the theorem for the finite case.

If the number of principal components of $Z-F$ is infinite, the sequence $\{Q_k\}$ is infinite. From § 7 we see that for a fixed k, $H\cdot K=\sum\alpha_{i_k}$, $i_k=1, 2, \cdots, k$, where each α_{i_k} is closed, none is void, and no two have common points. For $k+1$, $H\cdot K=\sum\alpha_{i_{k+1}}$, $i_{k+1}=1, 2, \cdots, k+1$, and each $\alpha_{i_{k+1}}$ is identical with or a part of some α_{i_k}. It follows that $H\cdot K$ is the sum of an infinite number of closed sets without common points, each one of which is the divisor of a monotone decreasing sequence $\alpha_{i_1}\supset\alpha_{i_2}\cdots$.

Let α and β be any two of these sets. For k sufficiently great, $\alpha\subset\alpha_{i_k}$, $\beta\subset\alpha_{i'_k}$, $i_k\neq i'_k$. By the first case H and K are irreducible between α_{i_k} and $\alpha_{i'_k}$. Hence, a fortiori, they are irreducible between α and β.

Remark. If the number of principal components of $Z-F$ is greater than two, both H and K are indecomposable, for then H and K are irreducible between each pair of three or more sets.

YALE UNIVERSITY, NEW HAVEN, CONN., U. S. A.

COLLINEATIONS OF PROJECTIVELY RELATED AFFINE CONNECTIONS.*

By M. S. Knebelman.

1. By a *collineation* of an n-dimensional affinely connected space we understand a point transformation which carries the paths of the space into paths. A collineation which preserves the affine parameter of the paths is called *affine* otherwise it is said to be *projective*.†

The following properties of finite continuous groups of collineations of an affinely connected space have been established by L. P. Eisenhart and the writer:‡ if Γ^i_{jk}, $i, j, k = 1, \cdots, n$, are the components of affine connection, B^i_{jkl} the components of the curvature tensor and ξ^i the components of an infinitesimal collineation of the group, then§

$$\xi^i{}_{,j,k} + \xi^l B^i_{jkl} = \delta^i_j \varphi_k + \delta^i_k \varphi_j, \tag{1.1}$$

where a subscript preceded by a comma denotes covariant differentiation with respect to the Γ's and φ_k are the components of a covariant vector given by

$$\varphi_k = \frac{1}{n+1}(\xi^h{}_{,h,k} + \xi^l S_{kl}). \tag{1.2}$$

A necessary and sufficient condition that a collineation be affine is that ξ^i satisfy the equations

$$\xi^i{}_{,j,k} + \xi^l B^i_{jkl} = 0, \tag{1.3}$$

that is, $\varphi_k = 0$.

The greatest number of parameters in a group of projective collineations is n^2+2n. Projectively flat spaces are the only spaces which admit a group of this number of parameters. The greatest number of parameters in a group of affine collineations is n^2+n. Flat spaces ($B^i_{jkl} = 0$) are the only spaces admitting a group of this number of parameters.

2. We propose to show how to reduce the problem of projective collineations to one of affine collineations, provided the number of parameters of the group, r, is $\leq n$. To this end we shall prove the following theorem:

If a given affine connection of components Γ^i_{jk} $(i, j, k = 1, \cdots, n)$ admits a finite continuous group G_r $(r \leq n)$ of collineations, projective or affine

* Received December 21, 1927.

† Knebelman, Proc. Nat. Acad. of Sc., vol. 13 (1927), pp. 396–400.

‡ Eisenhart & Knebelman, Proc. Nat. Acad. of Sc., vol. 13 (1927), pp. 38–42.

§ The summation convention of repeated indices is used throughout.

there exist projectively related affine connections for which G_r is a group of affine collineations. The determination of these projectively related connections depends upon n functions which are arbitrary in $n-r$ of the coördinates.

We first simplify our problem by means of a projective change of connection.* Two affine connections Γ and $\bar{\Gamma}$ are projectively related if and only if

$$\bar{\Gamma}^i_{jk} = \Gamma^i_{jk} + \delta^i_j \bar{\psi}_k + \delta^i_k \bar{\psi}_j, \tag{2.1}$$

where $\bar{\psi}_k$ are the components of an arbitrary covariant vector. For a given coördinate system we choose $\bar{\psi}_k = -\frac{1}{n+1}\bar{\Gamma}^h_{hk}$. Then, by contracting (2.1) for i and j we find $\bar{\Gamma}^h_{hk} = 0$. Since when a connection admits a finite continuous group of projective collineations all projectively related connections admit the same projective group,† there is no loss of generality in assuming $\Gamma^h_{hk} = 0$.

Let $\xi^i_{(\alpha)}$, $\alpha = 1, \cdots, r$, $(r \leqq n)$ indicate the components of the infinitesimal collineations of the group G_r that the affine connection is assumed to admit; then each vector must satisfy equations of the form (1.1), while (1.2) become, because of our assumption on the Γ's,

$$\varphi_{,k} = \frac{1}{n+1}\xi^h{}_{,h,k}. \tag{2.2}$$

$\xi^h{}_{,h,k}$ being the components of the gradient of the divergence of ξ^i.

In order to establish the first part of our theorem we must show that it is possible to choose a vector ψ_k so that the projectively related connection determined by it shall admit G_r as a group of affine collineations.

Equations (1.2) may be written‡ as

$$\begin{aligned} &\frac{1}{n+1}\left(\delta^i_j \frac{\partial^2 \xi^h}{\partial x^h \partial x^k} + \delta^i_k \frac{\partial^2 \xi^h}{\partial x^h \partial x^j}\right) \\ &= \frac{\partial^2 \xi^i}{\partial x^j \partial x^k} + \Gamma^i_{jh}\frac{\partial \xi^h}{\partial x^k} + \Gamma^i_{hk}\frac{\partial \xi^h}{\partial x^j} - \Gamma^h_{jk}\frac{\partial \xi^i}{\partial x^h} + \xi^h \frac{\partial \Gamma^i_{jk}}{\partial x^h} \end{aligned} \tag{2.3}$$

and the equations for the new connection must be

$$\begin{aligned} &\frac{\partial^2 \xi^i}{\partial x^j \partial x^k} + (\Gamma^i_{jh} + \delta^i_j \psi_h + \delta^i_h \psi_j)\frac{\partial \xi^h}{\partial x^k} + (\Gamma^i_{hk} + \delta^i_h \psi_k + \delta^i_k \psi_h)\frac{\partial \xi^h}{\partial x^j} \\ &- (\Gamma^h_{jk} + \delta^h_j \psi_k + \delta^h_k \psi_j)\frac{\partial \xi^i}{\partial x^h} + \xi^h\left(\frac{\partial \Gamma^i_{jk}}{\partial x^h} + \delta^i_j \frac{\partial \psi_k}{\partial x^h} + \delta^i_k \frac{\partial \psi_j}{\partial x^h}\right) = 0, \end{aligned} \tag{2.4}$$

if the theorem is to hold.

* Cf. T. Y. Thomas, Proc. Nat. Acad. of Sc. vol. 11 (1925), pp. 199–203.

† Loc. cit. p. 41.

‡ Loc. cit., p. 41.

When (2.3) and (2.4) are added, we find

$$\delta^i_j\left(\xi^h\frac{\partial\psi_k}{\partial x^h}+\psi_h\frac{\partial\xi^h}{\partial x^k}+\frac{1}{n+1}\frac{\partial^2\xi^h}{\partial x^h\,\partial x^k}\right)$$
$$+\delta^i_k\left(\xi^h\frac{\partial\psi_j}{\partial x^h}+\psi_h\frac{\partial\xi^h}{\partial x^j}+\frac{1}{n+1}\frac{\partial^2\xi^h}{\partial x^h\,\partial x^j}\right)=0,$$

which implies that the expression in each of the parentheses must vanish for each of the vectors $\xi^i_{(\alpha)}$. Therefore the nr equations that the n^2 derivatives $\partial\psi_i/\partial x^j$ must satisfy are

$$\xi^h_{(\alpha)}\frac{\partial\psi_i}{\partial x^h}+\psi_k\frac{\partial\xi^k_{(\alpha)}}{\partial x^i}+\frac{1}{n+1}\frac{\partial^2\xi^k_{(\alpha)}}{\partial x^k\,\partial x^i}=0. \tag{2.5}$$

We shall show that in virtue of the fact that the $\xi^i_{(\alpha)}$ form an r-parameter group, equations (2.5) are completely integrable. Since $\|\xi^i_\alpha\|$ is of rank r, we can arrange the components of the vectors in such a manner that $|\xi^\beta_{(\alpha)}|\neq 0$ $(\alpha,\beta=1,\cdots,r)$. Then the nr equations (2.5) can be solved for $\partial\psi_i/\partial x^\alpha$. To do this we introduce the r^2 quantities $\eta^{(\alpha)}_\beta$ defined by the equations

$$\eta^{(\alpha)}_\beta\,\xi^\beta_{(\gamma)}=\delta^\alpha_\gamma, \tag{2.6}$$

β being summed for $1,\cdots,r$. From these it follows that

$$\eta^{(\alpha)}_\beta\,\xi^\gamma_{(\alpha)}=\delta^\gamma_\beta. \tag{2.7}$$

We let

$$\frac{\partial\psi_i}{\partial x^\alpha}=F_{i\alpha},\qquad \frac{\partial\psi_i}{\partial x^\mu}=\psi_{i\mu}\qquad(\mu=r+1,\cdots,n),$$

so that the solution of (2.6) by Kramer's rule gives

$$F_{i\alpha}=-\psi_k\,\eta^{(\gamma)}_\alpha\frac{\partial\xi^k_{(\gamma)}}{\partial x^i}-\eta^{(\gamma)}_\alpha\,\xi^\mu_{(\gamma)}\,\psi_{i\mu}-\frac{1}{n+1}\eta^{(\gamma)}_\alpha\frac{\partial^2\xi^h_{(\gamma)}}{\partial x^h\,\partial x^i} \tag{2.8}$$
$$\begin{pmatrix} i,\,j,\,k\cdots=1,\cdots,n\\ \alpha,\beta,\gamma\cdots=1,\cdots,r\\ \lambda,\mu,\nu\cdots=r+1,\cdots,n\end{pmatrix}.$$

Equations (2.8) constitute a system of simultaneous partial differential equations studied by König.* We may write more generally

$$\frac{\partial\psi_i}{\partial x^\alpha}=F_{i\alpha}(x^j,\,\psi_k,\,\psi_{l\mu}). \tag{2.9}$$

* Cf. Math. Ann., vol. 23 (1884), pp. 520–526.

To obtain the conditions of complete integrability of (2.9) we regard x^j, ψ_k, $\psi_{l\mu}$ as independent variables in the partial differentiation of $F_{i\alpha}$. We thus get

$$\frac{\partial^2 \psi_i}{\partial x^\alpha \partial x^\beta} = \frac{\partial F_{i\alpha}}{\partial x^\beta} + \frac{\partial F_{i\alpha}}{\partial \psi_k}\frac{\partial \psi_k}{\partial x^\beta} + \frac{\partial F_{i\alpha}}{\partial \psi_{l\mu}}\frac{\partial \psi_{l\mu}}{\partial x^\beta}. \tag{2.10}$$

But

$$\frac{\partial \psi_k}{\partial x^\beta} = F_{k\beta}, \qquad \frac{\partial \psi_{l\mu}}{\partial x^\beta} = \frac{\partial F_{l\beta}}{\partial x^\mu} + \frac{\partial F_{l\beta}}{\partial \psi_k}\psi_{k\mu} + \frac{\partial F_{l\beta}}{\partial \psi_{k\lambda}}\frac{\partial \psi_{k\lambda}}{\partial x^\mu};$$

therefore, substituting these values into (2.10), interchanging α and β and subtracting, we obtain

$$\begin{aligned}\left(\frac{\partial F_{i\alpha}}{\partial x^\beta} - \frac{\partial F_{i\beta}}{\partial x^\alpha} + F_{h\beta}\frac{\partial F_{i\alpha}}{\partial \psi_h} - F_{h\alpha}\frac{\partial F_{i\beta}}{\partial \psi_h} + \frac{\partial F_{i\alpha}}{\partial \psi_{h\mu}}\frac{\partial F_{h\beta}}{\partial x^\mu} - \frac{\partial F_{i\beta}}{\partial \psi_{h\mu}}\frac{\partial F_{h\alpha}}{\partial x^\mu}\right)&\\ + \left(\frac{\partial F_{i\alpha}}{\partial \psi_{h\mu}}\frac{\partial F_{h\beta}}{\partial \psi_j} - \frac{\partial F_{i\beta}}{\partial \psi_{h\mu}}\frac{\partial F_{h\alpha}}{\partial \psi_j}\right)\psi_{j\mu}&\\ + \left(\frac{\partial F_{i\alpha}}{\partial \psi_{j\lambda}}\frac{\partial F_{j\beta}}{\partial \psi_{k\mu}} - \frac{\partial F_{i\beta}}{\partial \psi_{j\lambda}}\frac{\partial F_{j\alpha}}{\partial \psi_{k\mu}}\right)\frac{\partial \psi_{k\lambda}}{\partial x^\mu} = 0.&\end{aligned} \tag{2.11}$$

If (2.9) are to be completely integrable (2.11) must hold for arbitrary values of $\psi_{j\mu}$ and of $\frac{\partial \psi_{k\lambda}}{\partial x^\mu}$ $(= \partial^2 \psi_k / \partial x^\lambda \partial x^\mu)$. The first two parentheses of (2.11) do not contain $\partial \psi_{k\lambda}/\partial x^\mu$; we must therefore have

$$\left(\frac{\partial F_{ip}}{\partial x^q} + F_{hq}\frac{\partial F_{ip}}{\partial \psi_h} + \frac{\partial F_{ip}}{\partial \psi_{h\mu}}\frac{\partial F_{hq}}{\partial x^\mu} + \frac{\partial F_{ip}}{\partial \psi_{h\mu}}\frac{\partial F_{hq}}{\partial \psi_j}\psi_{j\mu}\right)\delta^{pq}_{\alpha\beta} = 0 \tag{2.12}$$

and

$$\delta^{pq}_{\alpha\beta}\left(\frac{\partial F_{ip}}{\partial \psi_{j\lambda}}\frac{\partial F_{jq}}{\partial \psi_{k\mu}} + \frac{\partial F_{ip}}{\partial \psi_{j\mu}}\frac{\partial F_{jq}}{\partial \psi_{k\lambda}}\right) = 0. \tag{2.13}$$

In order to verify (2.12) and (2.13) for our particular system (2.8) we make use of the relations between the generators and the constants of composition of the group.* Let these constants be denoted by $c^\gamma_{\delta\varepsilon}$; then we have

$$\xi^k_{(\delta)}\frac{\partial \xi^h_{(\varepsilon)}}{\partial x^k} - \xi^k_{(\varepsilon)}\frac{\partial \xi^h_{(\delta)}}{\partial x^k} = c^\gamma_{\delta\varepsilon}\,\xi^h_{(\gamma)} \qquad (\gamma, \delta, \varepsilon = 1, \cdots, r). \tag{2.14}$$

We define $n^2 r$ quantities $J^h_{k\alpha}$ by means of the equations

$$J^h_{k\alpha} = -\eta^{(\gamma)}_\alpha \frac{\partial \xi^h_{(\gamma)}}{\partial x^k}. \tag{2.15}$$

* Cf. L. P. Eisenhart, Riemannian Geometry, Chap. VI, in particular sec. 77.

When we multiply (2.14) by $\eta_\alpha^{(\delta)}\,\eta_\beta^{(\varepsilon)}$ and sum the δ's and the ε's we obtain, because of (2.7),

$$\eta_\alpha^{(\gamma)}\,\xi_{(\gamma)}^k\,J_{k\beta}^h - \eta_\beta^{(\gamma)}\,\xi_{(\gamma)}^k\,J_{k\alpha}^h = -\,c_{\delta\varepsilon}^{\gamma}\,\xi_{(\gamma)}^h\,\eta_\alpha^{(\delta)}\,\eta_\beta^{(\varepsilon)}. \tag{2.16}$$

From (2.15) we have, by means of (2.7),

$$\frac{\partial\,\xi_{(\gamma)}^h}{\partial\,x^k} = -\,J_{k\varepsilon}^h\,\xi_{(\gamma)}^{\varepsilon}, \qquad \frac{\partial\,\eta_\alpha^{(\gamma)}}{\partial\,x^i} = J_{i\varepsilon}^{\varepsilon}\,\eta_\varepsilon^{(\gamma)}, \tag{2.17}$$

ε being summed from 1 to r. From the first of (2.17) we have

$$\frac{\partial^2\,\xi_{(\gamma)}^h}{\partial\,x^i\,\partial\,x^j} = -\,\xi_{(\gamma)}^{\varepsilon}\,\frac{\partial\,J_{i\varepsilon}^h}{\partial\,x^j} - J_{i\varepsilon}^h\,\frac{\partial\,\xi_{(\gamma)}^{\varepsilon}}{\partial\,x^j} = \left(J_{i\varepsilon}^h\,J_{j\delta}^{\varepsilon} - \frac{\partial\,J_{i\delta}^h}{\partial\,x^j}\right)\xi_{(\gamma)}^{\delta}.$$

Interchanging i and j and subtracting the result from the above equation we obtain,

$$\frac{\partial\,J_{j\beta}^h}{\partial\,x^i} - \frac{\partial\,J_{i\beta}^h}{\partial\,x^j} = J_{j\varepsilon}^h\,J_{i\beta}^{\varepsilon} - J_{i\varepsilon}^h\,J_{j\beta}^{\varepsilon}. \tag{2.18}$$

When we differentiate (2.16) partially with respect to x^i, replace the partial derivatives of ξ and η from (2.17) and make use of (2.16) and (2.18) in the reduction we obtain

$$\eta_\alpha^{(\gamma)}\,\xi_{(\gamma)}^k\,\frac{\partial\,J_{i\beta}^h}{\partial\,x^k} - \eta_\beta^{(\gamma)}\,\xi_{(\gamma)}^k\,\frac{\partial\,J_{i\alpha}^h}{\partial\,x^k} + J_{i\beta}^k\,J_{k\alpha}^h - J_{i\alpha}^k\,J_{k\beta}^h = 0. \tag{2.19}$$

When we contract (2.19) for h and i we obtain

$$\eta_\alpha^{(\gamma)}\,\xi_{(\gamma)}^k\,\frac{\partial\,J_{h\beta}^h}{\partial\,x^k} - \eta_\beta^{(\gamma)}\,\xi_{(\gamma)}^k\,\frac{\partial\,J_{h\alpha}^h}{\partial\,x^k} = 0. \tag{2.20}$$

Equations (2.8), whose complete integrability we are trying to establish, may be written, because of (2.15), as

$$F_{i\alpha} = \psi_k\,J_{i\alpha}^k - \eta_\alpha^{(\gamma)}\,\xi_{(\gamma)}^{\mu}\,\psi_{i\mu} + \frac{1}{n+1}\left(\frac{\partial\,J_{h\alpha}^h}{\partial\,x^i} - J_{i\alpha}^{\gamma}\,J_{h\gamma}^h\right),$$

from which we get

$$\frac{\partial\,F_{i\alpha}}{\partial\,\psi_k} = J_{i\alpha}^k, \qquad \frac{\partial\,F_{i\alpha}}{\partial\,\psi_{j\lambda}} = -\,\delta_i^j\,\eta_\alpha^{(\gamma)}\,\xi_{(\gamma)}^{\lambda},$$

$$\frac{\partial\,F_{i\alpha}}{\partial\,x^l} = \psi_k\,\frac{\partial\,J_{i\alpha}^k}{\partial\,x^l} + \psi_{i\mu}\,(J_{l\alpha}^{\mu} - \eta_\delta^{(\gamma)}\,\xi_{(\gamma)}^{\mu}\,J_{l\alpha}^{\delta}) \tag{2.21}$$

$$+ \frac{1}{n+1}\left(\frac{\partial^2\,J_{h\alpha}^h}{\partial\,x^i\,\partial\,x^l} - J_{i\alpha}^{\gamma}\,\frac{\partial\,J_{h\gamma}^h}{\partial\,x^l} - J_h^h\,\frac{\partial\,J_{i\alpha}^{\gamma}}{\partial\,x^l}\right).$$

When these values are substituted into (2.13), the four terms cancel and (2.12) reduce to

$$\psi_k\left(J^h_{i\alpha}\,J^k_{h\beta}-J^h_{i\beta}\,J^k_{h\alpha}+\eta^{(\gamma)}_{\beta}\,\xi^h_{(\gamma)}\frac{\partial J^k_{i\alpha}}{\partial x^h}-\eta^{(\gamma)}_{\alpha}\,\xi^h_{(\gamma)}\frac{\partial J^k_{i\beta}}{\partial x^h}\right)$$

$$+\psi_{i\mu}\,\xi^{\mu}_{(\gamma)}\,\eta^{(\varepsilon)}_{\alpha}\,\eta^{(\delta)}_{\beta}\,(c^{\gamma}_{\varepsilon\delta}+c^{\gamma}_{\delta\varepsilon})-\frac{1}{n+1}\,\frac{\partial}{\partial x^i}\left(\eta^{(\gamma)}_{\alpha}\,\xi^k_{(\gamma)}\,\frac{\partial J^h_{h\beta}}{\partial x^k}-\eta^{(\gamma)}_{\beta}\,\xi^k_{(\gamma)}\,\frac{\partial J^h_{h\alpha}}{\partial x^k}\right),$$

and this vanishes because of (2.19), (2.20) and the fact that the constants of composition of a group are skew-symmetric. Therefore equations (2.8) are completely integrable and their general solution is a set of n functions ψ_k such that for $x=x_0$ they become arbitrarily chosen functions of $x^{r+1},\cdots,x^n$. Thus our theorem is completely proved.*

As a corollary of the above theorem we have when $r=n$:

If an affine connection admits a simply transitive group G_n of collineations there exist projectively related affine connections admitting G_n as a group of affine collineations; the determination of these projectively related connections depends upon n arbitrary constants.

---

* For $r=1$ the proof is due to L. P. Eisenhart, cf. sec. 46 Non-Riemannian Geometry. His method of normalizing the vector is not suitable for the general case.

PRINCETON UNIVERSITY.

THE DISCRIMINANT OF THE m-ARY QUADRATIC IN THE GALOIS FIELDS OF ORDER 2^n.*

By A. D. Campbell.

Let us consider the general m-ary quadratic in the Galois fields† of order 2^n

$$f(x_1, x_2, \cdots, x_m) \equiv \sum a_{ij} x_i x_j = 0, \tag{1}$$

where

$$i, j = 1, 2, \cdots, m; \qquad j \geqq i; \qquad a_{ji} = 0 \quad \text{if} \quad j \neq i.$$

We call the quadratic (1) degenerate if it can be reduced by a non-singular projective transformation to an equation involving less than m variables. Thus $\sum a_{ii} x_i^2 = 0$ is reducible to $x_1'^2 = 0$. We call the discriminant Δ of (1) the algebraic expression that connects the coefficients of this quadratic and whose vanishing is the necessary and sufficient condition that (1) be degenerate. In this paper we derive the discriminant of (1). The presence of the modulus 2 in these $GF(2^n)$ causes the discussion of the discriminant‡ that is used in the real or complex algebra to break down here. We define the derivative dx_i^r/dx_i as rx_i^{r-1}. In algebraic operations with the numbers of any Galois field of order p^n we can proceed as though the numbers involved were not reducible modulo p or by any other modulus, and then to obtain our final results replace p and any other modulus by zero in the algebraic expressions. Thus $(x+y)^3 = x^3+3x^2y+3xy^2+y^3$ in the $GF(2^n)$ as well as elsewhere; but $3 \equiv 1 \pmod 2$ hence we write $(x+y)^3 = x^3+x^2y+xy^2+y^3$ in the $GF(2^n)$.

We define the polar of any point $P'(x_1', x_2', \cdots, x_m')$ with respect to (1) as

$$\sum \frac{\partial f}{\partial x_i} x_i' = 0, \qquad (i = 1, 2, \cdots, m). \tag{2}$$

If the polar (2) is to pass through a fixed point $P''(x_1'', x_2'', \cdots, x_m'')$ for every point P' we must have

$$\frac{\partial f}{\partial x_1''} = 0, \quad \frac{\partial f}{\partial x_2''} = 0, \cdots, \frac{\partial f}{\partial x_m''} = 0, \tag{3}$$

* Received October 17, 1927, in revised form January 9, 1928.

† We call these fields more briefly $GF(2^n)$.

‡ Compare Bôcher, Introduction to Higher Algebra, p. 128.

where $\partial f/\partial x_i''$ means the partial derivative $\partial f/\partial x_i$ with $x_1'', x_2'', \cdots, x_m''$ substituted for $x_1, x_2, \cdots, x_m$ respectively. For (3) to have a solution $x_1'', x_2'', \cdots, x_m''$ where the x''s are not all zero the determinant D of the following matrix M must vanish

$$(4) \qquad M \equiv \begin{Vmatrix} 2a_{11}, & a_{12}, & \cdots, & a_{1m} \\ a_{12}, & 2a_{22}, & \cdots, & a_{2m} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{1m}, & a_{2m}, & \cdots, & 2a_{mm} \end{Vmatrix}.$$

We do not immediately put $2 \equiv 0$ in M or in D but we wait to do this until later in the discussion. We shall show that the discriminant Δ of (1) is obtained from D. First we prove the relative invariance of M in (4) under any non-singular projective transformation T with determinant $E \neq 0$. In Dickson "Linear Groups" § 100 we see that T can be expressed as a product BC_m where B is a product of transformations of the form

$$(5) \qquad x_r = x_r' + \lambda x_s', \quad x_i = x_i', \qquad (i = 1, 2, \cdots, m;\ i \neq r,\ r \neq s),$$

with λ arbitrary, and where C_m denotes the transformation

$$(6) \qquad x_i = x_i' \ (i = 1, 2, \cdots, m-1), \qquad x_m = \tau x_m', \ (\tau = E).$$

We note that the determinant of (5) is unity while that of (6) is E.

The transformation (5) sends the matrix M of (4) into

$$(4') \qquad M' \equiv \begin{Vmatrix} 2a_{11}, & a_{12}, & \cdots, & a_{1s}+\lambda a_{1r}, & \cdots, & a_{1m} \\ a_{12}, & 2a_{22}, & \cdots, & a_{2s}+\lambda a_{2r}, & \cdots, & a_{2m} \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ a_{1s}+\lambda a_{1r}, & a_{2s}+\lambda a_{2r}, & \cdots, & 2a_{ss}+2\lambda a_{rr}, & \cdots, & a_{sm}+\lambda a_{rm} \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ a_{1m}, & a_{2m}, & \cdots, & a_{sm}+\lambda a_{rm}, & \cdots, & 2a_{mm} \end{Vmatrix}.$$

The matrix M' is readily seen to be a product of matrices of the form $L'ML''$ where L' has all its elements zero except for those of the main diagonal which are all unity and for the element in the sth row and rth column which is λ, L'' is like L' except that λ occurs in the rth row and sth column and not as in L'. The determinants of L' and L'' are both unity. The transformation (6) sends M into

$$(4'') \qquad M'' \equiv \begin{Vmatrix} 2a_{11}, & a_{12}, & \cdots, & \tau a_{1m} \\ a_{12}, & 2a_{22}, & \cdots, & \tau a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \tau a_{1m}, & \tau a_{2m}, & \cdots, & 2\tau^2 a_{mm} \end{Vmatrix},$$

which matrix is the product FMF where F is the matrix of (6). The product BC_m therefore sends M into L_1FMFL_2, where L_1 is a product of matrices of the form L' and L_2 a product of matrices of the form L'', and therefore L_1 and L_2 have determinants each equal to unity. Thus we have shown that BC_m sends the determinant D of M in (4) into $D' \equiv E^2D$ where E is the determinant of the non-singular transformation T with which we started. Therefore the necessary and sufficient condition for D' (or D) to be zero is that D (or D') be zero.

In Dickson, loc. cit., § 199 we find that if (1) is nondegenerate it can be put by a transformation T in one of the forms

$$(7) \quad x_1x_2 + x_3x_4 + \cdots + x_{m-2}x_{m-1} + x_m^2 = 0 \ (m \text{ odd}),$$

$$(8) \quad x_1x_2 + x_3x_4 + \cdots + x_{m-3}x_{m-2} + x_{m-1}x_m + \alpha x_{m-1}^2 + \alpha x_m^2 = 0,$$

where α is zero or is a particular one of the values α' for which $x_{m-1}x_m + \alpha' x_{m-1}^2 + \alpha' x_m^2$ is irreducible in the $GF(2^n)$. If (1) is degenerate it can be reduced by the methods of Dickson, loc. cit. § 199, to (7) or (8) with $m' < m$ variables. Such a quadratic in m' variables when considered as a quadratic in m variables certainly is such that the corresponding determinant D' vanishes. But $D' \equiv E^2D$; hence if (1) is degenerate then $D = 0$.

We note that, if m is even in (1) and (1) is nondegenerate and we reduce (1) to the form (8), then the determinant $D' \equiv E^2D$ of (8) does not vanish, so $D \neq 0$. Hence a necessary and sufficient condition for (1) with m even to be degenerate is the vanishing of the determinant D of M in (4). Putting $2 \equiv 0$ in this determinant we get the discriminant of (1) for m even, namely

$$(9) \qquad \Delta \equiv \begin{Vmatrix} 0, & a_{12}, & a_{13}, & \cdots, & a_{1m} \\ a_{12}, & 0, & a_{23}, & \cdots, & a_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{1m}, & a_{2m}, & a_{3m}, & \cdots, & 0 \end{Vmatrix}.$$

On the other hand, if m is odd in (1) and (1) is non-degenerate and we reduce (1) to the form (7), then the determinant $D' \equiv E^2D$ of (7)

vanishes, so $D = 0$. But $D' \equiv 2 \equiv 0 \pmod 2$, whereas $E^2 \not\equiv 0 \pmod 2$ so E^2 cannot have 2 as a factor; therefore we must have $D \equiv 2G$. From the identity $D' \equiv 2 \cdot 1 \equiv E^2 \cdot 2 \cdot G$ we see that $E^2 G \equiv 1$, so G cannot have 2 as a factor. If (1) is degenerate, with m odd, and is reduced to (7) or (8) with $m' < m$, then from the identity $D' \equiv E^2 D \equiv E^2 \cdot 2G \equiv 0$ (not modulo 2) we see that we have $G = 0$. Hence the necessary and sufficient condition for (1) to be degenerate when m is odd is that G shall vanish in $D \equiv 2G$. Therefore the discriminant Δ of (1) for m odd is obtained by expanding the determinant $D \equiv 2G$ and from this expanded form removing the factor 2 and finally putting $2 \equiv 0$ in the remaining part of D, i. e. $\Delta \equiv G$ (where we have put $2 \equiv 0$ in G).

CONCERNING THE COMPLEMENTARY DOMAINS OF CONTINUA.*

By Gordon T. Whyburn.

1. Introductory. Definitions and notations. A point set M is said to be connected in the weak sense, or merely *connected,* provided it is not the sum of any two mutually separated point sets. A set M is said to be connected in the strong† sense, or strongly connected, provided that every two points of M lie together in some continuum which is a subset of M. A collection G of point sets will be said to have property E provided that for any $\varepsilon > 0$ not more than a finite number of elements of G are of diameter $> \varepsilon$. If the collection of all the complementary domains of a continuum M has property E, then M will be called an *„E-continuum".*

In this paper questions relating to the connectivity in both of the above senses of the sum of the boundaries of the complementary domains of various classes of continua will be considered. It is to be understood that the point sets considered lie in a Euclidean space of two dimensions, although some of the theorems, as indicated below, are true in n-dimensional space. The following notations will be used:

$\varrho(X, Y)$ for the distance from X to Y,

$\delta(M)$ for the diameter of M,

$\supset$ for „contains",

$\subset$ for „is contained in", or „is a subset of",

$\overline{M}$ for $M + M'$, i. e., for the set M plus all its limit points,

$F(R)$ for the boundary of R, where R is a domain,

$K \cdot H$ for the set of points common to the sets K and H.

2. Continua in general. In this section a study will be made of E-continua. The point P will be said to be *locally isolated by continua* of a class X provided that for every $\varepsilon > 0$ a continuum N of the class X exists such that $\delta(N) < \varepsilon$ and P belongs to a bounded complementary domain of N.

THEOREM 1. *Every point P of an E-continuum M which belongs to the boundary of no complementary domain of M is locally isolated by subcontinua of M.*

Proof. Let d be any number > 0 and C_1 a circle with center P and diameter $d/4$. Since M is an E-continuum and P is a boundary point of

* Received August 23, 1927. Presented to the American Mathematical Society, February 26, April 15, and September 9, 1927.

† Cf. R. L. Moore, Proc. Nat. Acad. Sci., vol. 9 (1923), pp. 101–106.

no complementary domain of M, it follows that there exists within C_1 a circle C_2 with center P which neither contains nor encloses a point of any complementary domain of M which contains a point of C_1. There exists a domain D such that $F(D) = C_1 + C_2$. Let $K = M \cdot (C_1 + C_2 + D)$. Then K is closed and $\delta(K) < d$. Let R_p denote the complementary domain of K which contains P. Then R_p lies wholly within C_1. For if, on the contrary, it contains a point Q on or without C_1, then $R_p \supset PQ$, an arc from P to Q, and $PQ \supset AB$, an arc such that $A \subset C_1$, $B \subset C_2$, $AB \subset (C_1 + C_2 + D)$. But then AB belongs to some complementary domain of M, contrary to the fact that no such domain contains points of both C_1 and C_2. Therefore R_p lies wholly within C_1.

Now let N denote the boundary of the unbounded complementary domain of $F(R_p)$. By a theorem of Chittenden's,* N is a continuum. As $N \subset K$, then $\delta(N) < d$. Clearly P lies in a bounded complementary domain of N. Hence P is locally isolated by subcontinua of M, and our theorem is proved.

THEOREM 2. *Let H denote the sum of the boundaries of all the complementary domains of an E-continuum M. Then M is connected im kleinen at every point of $M - H$.*

Theorem 2 is an immediate consequence of Theorem 1 and the definition of connectedness im kleinen.

THEOREM 3. *In order that a bounded continuum M should be an E-continuum it is necessary and sufficient that every point of M should be locally isolated by continua each of which is a subset of M plus a finite number of its complementary domains.*

Proof. The condition is necessary. For let P be any point of a bounded, E-continuum M. If P belongs to the boundary of no complementary domain of M, then by Theorem 1, P is locally isolated by subcontinua of M. So let us suppose that P belongs to the boundary of at least one complementary domain of M. Let d be any positive number and C_1 a circle with center P and diameter $d/2$. Since M is an E-continuum, there exists within C_1 a circle C_2 with center P such that if g is any complementary domain of M which contains a point of C_1 then either g contains no point on or within C_2 or else P belongs to the boundary of g. There are not more than a finite number of complementary domains g of M such that g contains a point of C_1 and also has P for a boundary point. Let T denote the sum of all such domains which have this property. Then $M + T$ is closed and is a subset of M plus a finite number of its complementary domains. There exists a domain D such that $C_1 + C_2 = F(D)$. Let $K = (M + T) \cdot (C_1 + C_2 + D)$. Then K is closed and $\delta(K) < d$; and by almost

* Bull. Amer. Math. Soc., vol. 28 (1922), pp. 310–312.

the same argument as in the proof of Theorem 1 it is shown that K contains a continuum N such that P lies in a bounded complementary domain of M. Thus the condition of Theorem 3 is necessary.

The condition is also sufficient. For suppose a bounded continuum M satisfies the condition, but suppose, contrary to this theorem, that for some $\varepsilon > 0$ there exists an infinite sequence of complementary domains $D_1, D_2, \cdots$ of M, with $\delta(D_n) > \varepsilon$ for every n. Then for every n, D_n contains two points X_n and Y_n with $\varrho(X_n, Y_n) > \varepsilon$. The set of points $X_1 + X_2 + X_3 + \cdots$ has at least one limit point P which, it is easy to see, must belong to M. By hypothesis there exists a continuum N with $\delta(N) < \varepsilon/30$ such that $N \subset M +$ a finite number of the complementary domains of M and such that P lies in a bounded complementary domain R of N. There exists an infinite sequence of positive integers $n_1, n_2, \cdots$, such that, for every i, $X_{n_i} \subset R$ and $\varrho(Y_{n_i}, P) > \varepsilon/2$. Hence for no i does Y_{n_i} belong to R; therefore, for each i, N contains at least one point of D_{n_i}. But this is contrary to the fact that N contains points of at most a finite number of the complementary domains of M. Hence the condition of Theorem 3 is sufficient.

Theorems 1, 2, and 3 are true in n-dimensions since it is true* in n-dimensions that the common boundary of two domains is connected.

THEOREM 4. *Every indecomposable E-continuum M is the boundary of at least one of its complementary domains.*

Proof. Suppose, on the contrary, that M is not the boundary of any one of its complementary domains. Then if $B_1, B_2, \cdots$ denote the boundaries of the complementary domains of M, for no integer i is $B_i \equiv M$. Hence, for each i, B_i is a proper subcontinuum of M. Therefore, $M \neq \sum_{i=1,2,3,\cdots} B_i$, because by a theorem due to Janiszewski and Kuratowski,† no indecomposable continuum is the sum of any countable collection of its proper subcontinua. Hence there exists a point P of M which belongs to none of the sets B_i. But since M is an E-continuum, by Theorem 2, M is connected im kleinen at P. This is impossible, because M is indecomposable. This contradiction proves Theorem 4.

Theorem 4 is a generalization of a result of Urysohn's‡ that every indecomposable continuum which has only a finite number of complementary domains is the boundary of at least one of them. Urysohn gives an example showing the existence of an indecomposable continuum which is not the boundary of any one of its complementary domains. However, it

* For this theorem see P. Urysohn, Fundamenta Math., vol. 7 (1925), p. 95.

† Fundamenta Math., vol. 1 (1920), p. 214.

‡ P. Urysohn, loc. cit.

is true that *every proper subcontinuum of any indecomposable continuum M is the boundary of some domain.* For let K be any proper subcontinuum of an indecomposable continuum M. By a theorem of R.L. Moore's* $M-K$ is connected. Hence if R denotes the complementary domain of K containing $M-K$, then since every point of K is a limit point of $M-K$ it follows that $F(R) = K$.

THEOREM 5. *In order that the sum H of the boundaries of all the complementary domains of a bounded E-continuum M should be connected in the strong sense it is necessary and sufficient that $S-(M-H)$ should be connected in the strong sense.*

Proof. The condition is sufficient. For let M be a continuum satisfying all the conditions of this theorem, and let A and B be any two points of H. By hypothesis $S-(M-H)$ contains a continuum N containing both A and B. Let K denote the point set obtained by adding together the boundaries of all those complementary domains of M whose boundary contains at least one point of N. Clearly K contains both A and B and is a subset of H. It remains to show that K is a continuum.

The set K is closed. For since M is an E-continuum it is easy to see that no point which belongs to neither N nor K can be a limit point of K; and any limit point of K which belongs to N must belong also to M and hence must belong to H and to K. Therefore K is closed. Now let us suppose, contrary to what we purpose to show, that K is not connected. Then by a theorem due to Knaster and Kuratowski† there exist points p and q of K and a bounded continuum L such that $K \cdot L = 0$ and p and q lie in different complementary domains of R_p and R_q respectively of L. Each of the sets M and N must contain a point in each of the domains R_p and R_q. And since M and N are connected, they must contain points P_M and P_N respectively in common with L. The point P_N cannot belong to M, for $M \cdot N \subset H \cdot N \subset K$, and $K \cdot L = 0$. Hence P_N belongs to some complementary domain R of M. And since L is not a subset of R, therefore it must contain a point of $F(R)$. But $F(R) \subset K$, and $K \cdot L = 0$. Thus the supposition that K is not connected leads to a contradiction. Therefore K is a continuum, and H is strongly connected.

The condition is also necessary. For let M be any continuum in a plane S such that the sum H of the boundaries of all the complementary domains of M is strongly connected, and let A and B be any two points of $S-(M-H)$. Let A_0 and B_0 denote A and B, respectively, or points belonging to the boundaries of the complementary domains R_a and R_b of M which contain A and B respectively according as A and B, or either A or B,

* Proc. Nat. Acad. Sci., vol. 12 (1926), pp. 359–363.

† Fundamenta Math., vol. 2 (1921), p. 233, Theorem 37.

do or do not belong to M. By hypothesis there exists a continuum N which contains $A_0 + B_0$ and is a subset of H. Then clearly the continuum $N + \bar{R}_a + \bar{R}_b$ contains $A + B$ and is a subset of $S - (M - H)$. Therefore $S - (M - H)$ is strongly connected, and our theorem is proved.

The following example shows that the condition of Theorem 5 is not sufficient in the absence of the restriction that the collection of all the complementary domains of the continuum in question shall have property E. Let J denote the simple closed curve formed by joining by straight line intervals in the order here named the points $(0, \frac{1}{8})$, $(-\frac{1}{8}, \frac{1}{8})$, $(-\frac{1}{8}, 1\frac{1}{8})$, $(2\frac{1}{8}, 1\frac{1}{8})$, $(2\frac{1}{8}, -\frac{1}{8})$, $(-\frac{1}{8}, -\frac{1}{8})$, $(-\frac{1}{8}, \frac{1}{8})$, $(0, \frac{1}{8})$, $(0, 0)$, $(2, 0)$, $(2, 1)$, $(0, 1)$, and $(0, \frac{1}{8})$. Let R denote J plus its interior. Let L denote the interval of the Y axis from $(0, 0)$ to $(0, 1)$, and for each $n = 0, 1, 2, \cdots$ let L_n be defined by the relations $x = 1/2^n$, $0 \leqq y \leqq 1$. For each $n = 0, 1, 2, 3, \cdots$, let R_{1n} be a rectangle, plus its interior, whose base lies on the interval $(0, 0)$ to $(1, 0)$ of the X axis and contains the point $(1/2^n, 0)$ and is of length $1/10^{n+1}$, and whose altitude is 1/8 unit long; and let R_{2n} be a rectangle, plus its interior, whose base lies on the line $y = 1$ and contains the point $(1/2^n, 1)$ and of length $1/10^{n+1}$, and whose altitude is 1/8 unit long. Then if M denotes the continuum

$$R + L + \sum_{n=0}^{\infty} [L_n + R_{1n} + R_{2n}],$$

and H denotes the sum of the boundaries of all the complementary domains of M, then $S - (M - H)$ is strongly connected, but the points $(2, 1)$ and $(0, 1/2)$ of H cannot be joined by any continuum which is a subset of H.

3. **Continuous curves in general.** THEOREM 6. *In order that the sum H of the boundaries of all the complementary domains of a continuous curve M in a plane S should be connected in the strong sense it is necessary and sufficient that $S - (M - H)$ should be connected in the strong sense.*

Theorem 6 is an immediate consequence of Theorem 5 and Schoenfliess' theorem* that every continuous curve is an E-continuum.

THEOREM 7. *In order that the sum H of the boundaries of all the complementary domains of a continuous curve M, or indeed, of any bounded continuum M,† in a plane S should be connected it is necessary and sufficient that $S - (M - H)$ should be connected.*

Proof. The condition is sufficient. For suppose $S - (M - H)$ is connected, but suppose, contrary to this theorem, that H itself is not connected. Then by Knaster and Kuratowski's‡ theorem there exist two points A and B

* A. Schoenfliess, *Die Entwicklung der Lehre von den Punktmannigfaltigkeiten*, Zweiter Teil, Leipzig, 1908, p. 237.

† That this theorem could be so extended was suggested to me by Professor R. L. Moore.

‡ Loc. cit.

of H and a bounded continuum N such that $N \cdot H = 0$ and A and B lie in different complementary domains of N. Since both M and $S-(M-H)$ are connected, and each contains the points A and B, therefore they contain points P_m and P_r respectively in common with N. Since $N \cdot H = 0$, therefore P_r cannot belong to M, and hence it belongs to some complementary domain R of M. But N is not a subset of R, for $N \supset P_m \subset M$. Therefore $F(R)$ contains at least one point P of N. But $F(R) \subset H$; hence $P \subset H$, contrary to the fact that $N \cdot H = 0$. Thus the supposition that H is not connected leads to a contradiction.

The condition is also necessary. For suppose H is connected. Let T denote complement of M, i. e., $T = S - M$. Then $S-(M-H) = H+T$; and since H contains at least one limit point of each maximal connected subset of T, therefore $H+T = S-(M-H)$ is connected. This completes the proof.

The following example shows that the sum of the boundaries of all the complementary domains of a plane continuous curve may be connected in the weak sense but not connected in the strong sense. Let the point set R be defined exactly as in the example given above at the end of § 2. Let I denote the interval of the Y axis from (0, 0) to (0, 1). Let S_1 and S_2 be two wavy curves similar in shape to the curve $y = \sin 1/x$, and such that (1) S_1 and S_2 both lie within the rectangle whose vertices are (0, 0), (0, 1), (2, 1), and (2, 0), (2) S_1 and S_2 have in common exactly one point P, and (3) every point of I is a limit point both of S_1 and S_2 but S_1+I and S_2+I are continua. Let D denote the bounded domain which is complementary to S_1+S_2+I. Let M denote the continuous curve obtained by adding to $R+I+S_1+S_2+D$ enough horizontal "cross-pieces" to make M connected im kleinen at every point of I, and let H denote the sum of the boundaries of all the complementary domains of M. Then H is connected, but clearly the points (2, 0) and (0, 1/2) of H lie together in no continuum which is a subset of H. Hence H is not strongly connected. It is to be observed that in this example the continuous curve M contains a subcontinuum which is not a continuous curve, as it necessarily must in view of the results in § 5 below.

THEOREM 8. *Let H denote the sum of the boundaries of all the complementary domains of a continuous curve M, or, indeed, of any bounded continuum M, and suppose that $M-H$ contains no continuum. Then if K is any connected open* subset of M, the set of points $K \cdot H$ is connected.*

Proof. Let W denote the set of points $K \cdot H$, and suppose, contrary to this

* The subset K of a point set M is said to be an open subset of M provided that the set of points $M-K$ is either vacuous or closed. The set K is said to be open relative to M provided that K contains no limit point of $M-K$.

theorem, that W is not connected. Then by Knaster and Kuratowski's theorem,* there exist two points A and B of W and a bounded continuum N such that $N \cdot W = 0$ and A and B lie in different complementary domains of W. Now since K is connected and contains both A and B, therefore it contains at least one point Y of N. By a theorem of Janiszewski's† it follows that N contains a subcontinuum F containing Y but such that $F \cdot (M - K) = 0$. Now F can contain no point of H, because, being a subset of N, it can contain no point of W; and as F contains no point of $M - K$ and $H - W$ is a subset of $M - K$, therefore F contains no point of $H - W$ and hence contains no point whatever of H. Since by hypothesis the set of points $M - H$ contains no continuum, therefore F contains a point X not belonging to $M - H$. The point X does not belong to H, for $F \cdot H = 0$. Therefore X belongs to some complementary domain R of M. Since F is connected and contains the point X of R and the point Y not belonging to R, therefore it contains a point Z of $F(R)$. But then Z belongs to H and, as shown above, $F \cdot H = 0$. Thus the supposition that W is not connected leads to a contradiction.

THEOREM 9. *Let H denote the sum of the boundaries of all the complementary domains of a continuous curve M. Then if $M - H$ contains no continuum, H is connected im kleinen.*

Proof. Let P be any point of H and C any circle with center at P. Let K denote the maximal connected subset of M which contains P and lies within C. Then K is a connected open subset of M. And since by hypothesis the set of points $M - H$ contains no continuum, therefore by Theorem 8, the set of points $H \cdot K$ is connected. And since P is not a limit point of $H - H \cdot K$, for $(H - H \cdot K) \subset (M - K)$, there exists a circle C_1 with center at P which encloses no point of $H - H \cdot K$. Then any point of H lying within C_1 lies together with P in the connected subset $H \cdot K$ of H which lies within C. Hence H is connected im kleinen at P, and the theorem is proved.

The following example shows that Theorem 9 is not true, even though H is arcwise connected, if the condition that the set $M - H$ shall contain no continuum be omitted. Let R denote the rectangle whose vertices are (0, 0), (0, 1), (2,1), und (2,0). For each positive integer n, let L_n be defined by the relations $x = 1/2^{n-1}$, $0 \leqq y \leqq 1$. For each odd positive integer n, let R_n be a rectangle, plus its interior, whose base lies on the X axis and contains the point $(1/2^{n-1}, 0)$ and is of length $1/10^n$, and whose altitude is of length 3/4. For each even positive integer n, let R_n be a similar rectangle plus its interior, except that its base lies on the

* Knaster and Kuratowski, loc. cit.

† Journ. de l'Ec. Pol., (2), vol. 16 (1912), p. 109.

line $y = 1$ and contains the point $(1/2^{n-1}, 1)$. For each positive integer n, let I_n be the sum of n intervals cut off between L_n and the Y axis on n lines parallel to the X axis and equally spaced between the X axis and the line $y = 1$. Let M denote the continuous curve

$$R + \sum_{n=1}^{\infty} [L_n + R_n + I_n].$$

Then if H denotes the sum of the boundaries of all the complementary domains of M, H is arcwise connected; but clearly it is not connected im kleinen at any point of the segment of the Y axis from (0,1/4) to (0, 3/4).

4. Continuous curves all of whose subcontinua are continuous curves.

THEOREM 10. *If H denotes the sum of the boundaries of all the complementary domains of a continuous curve M every subcontinuum of which is a continuous curve, then the set of points $M - H$ contains no continuum.*

Proof. Suppose, on the contrary, that $M - H$ contains a continuum T consisting of more than one point. Then since no point of T belongs to the boundary of any complementary domain of M, by a theorem of the author's* it follows that for each point X of T and for each positive integer n, M contains a simple closed curve J_{xn} which encloses X and is of diameter less than $1/n$. For each positive integer n, let G_n denote the point set obtained by adding together the interiors of all the simple closed curves of the collection $\{J_{xn}\}$. Since T is connected, it follows that for each n, G_n is a connected and bounded domain which contains T. Then for each n, the boundary B_n of the unbounded complementary domain of the boundary of G_n is a continuum. Clearly B_n is a subset of M, and it is of diameter $> d$, the diameter of T, because T lies in one of the bounded complementary domains of B_n. Let d_1 denote the shortest distance between the two continua B_1 and T. There exists a positive integer n_1 such that $1/n_1 < d_1/2$. Then it is easy to see that the continua B_1 and B_{n_1} can have no point in common. Let d_2 denote the shortest distance between the continua B_{n_1} and T. There exists a positive integer n_2 such that $1/n_2 < d_2/2$. Then B_{n_2} can have no point in common with either B_1 or B_{n_1}. This process may be continued indefinitely, and thus we obtain an infinite sequence of continua $B_1, B_{n_1}, B_{n_2}, B_{n_3}, \cdots$, each of which is of diameter greater than d and no two of which have a point in common. But since each of these continua belongs to M, this contradicts H. M. Gehman's theorem† that a continuous curve every subcontinuum of which is a continuous curve cannot contain more than a finite number of mutually exclusive

* G. T. Whyburn, Trans. Amer. Math. Soc., vol. 29 (1927), pp. 369–400.

† H. M. Gehman, Annals of Math., (2), vol. 27 (1925), pp. 29–46.

continua each of diameter greater than some preassigned positive number. Thus the supposition that $M-H$ contains a continuum leads to a contradiction.

THEOREM 11. *The sum H of the boundaries of all the complementary domains of a continuous curve M every subcontinuum of which is a continuous curve is connected and connected im kleinen.*

Proof. The set of points $M-H$ does not disconnect the plane, because if it did it would contain* a continuum which separates the plane; and by Theorem 10, $M-H$ contains no continuum whatever. Therefore, if S denotes the plane, the set of points $S-(M-H)$ is connected. Then by Theorem 7 it follows that H is connected. That H is connected im kleinen follows at once from Theorems 10 and 9.

5. In this section it will be shown that if H denotes the sum of the boundaries of all the complementary domains of a continuous curve M every subcontinuum of which is a continuous curve, then H is arcwise connected and $M-H$ is totally disconnected.

Definitions. The limit point P of a point set R is said to be *regularly accessible*† from R provided that for every positive number ϵ a $\delta_{\epsilon p}>0$ exists such that every point X of R with $\varrho(X, P)<\delta_{\epsilon p}$ can be joined to P by an arc XP with $\delta(XP)<\epsilon$ and $XP-P \subset R$. The point set M is said to have property S‡ provided that for each $\epsilon>0$, M is the sum of a finite number of connected point sets each of diameter $<\epsilon$. A point set M is said to be arcwise connected im kleinen provided that if P is any point of M, and $\epsilon>0$, a $\delta_{\epsilon p}>0$ exists such that each point X of M such that $\varrho(X, P)<\delta_{\epsilon p}$ can be joined in M to P by an arc XP with $\delta(XP)<\epsilon$.

THEOREM 12. *If R is any connected and relatively open (i.e., relative to M) subset of an arcwise connected im kleinen point set M, then R is arcwise connected.*

Proof. Since R is open relative to M, and M is arcwise connected im kleinen, it readily follows that R is arcwise connected im kleinen. Now let A and B be any two points of R, and let S_a denote A plus all those points of R which can be joined to A by an arc in R. Let S_b denote the set of points $R-S_a$. Then unless our theorem is true, the set S_b exists and contains the point B. But since R is connected, one of the sets S_a and S_b must contain a limit point of the other. Suppose S_a contains a points P which is a limit point of S_b. Then since R is arcwise connected im kleinen, it is easy to see that some point X of S_b can be joined in R to P by an arc XP; and since P belongs to S_a, it can be joined in R to A by an arc AP. But then the sum of the arcs XP and AP contains an arc AX from A to X which lies in R. This is con-

* Cf. Knaster and Kuratowski, loc. cit.

† Cf. G. T. Whyburn, Proc. Nat. Acad. Sci., vol. 13 (1927), pp. 650–657.

‡ R. L. Moore, Fundamenta Math., vol. 3 (1922), pp. 232–237.

trary to the fact that X belongs to S_b. A similar contradiction is obtained if we assume that some point P of S_b is a limit point of S_a. Thus the supposition that the set S_b exists leads to a contradiction, and it follows that R is arcwise connected.

THEOREM 13. *In order that every point of the boundary of a connected, bounded, and arcwise connected im kleinen point set R should be regularly accessible from R it is necessary and sufficient that R should have property S.*

Theorem 13 is a generalization of Theorem 1 of my paper *Concerning the open subsets of a plane continuous curve* (loc. cit.), and it may be proved by exactly the same method as used to prove that theorem, using the fact established in Theorem 12 that every connected open (relative to R) subset of R is arcwise connected.

THEOREM 14. *Every arcwise connected subset of a continuous curve every subcontinuum of which is a continuous curve has property S.*

Proof. Let R be any arcwise connected subset of a continuous curve M every subcontinuum of which is a continuous curve, and suppose, contrary to this theorem, that R does not have property S. Then it readily follows* that there exists a positive number ϵ such that R contains an infinite sequence of points $P_1, P_2, P_3, \cdots$, having a sequential limit point P and such that no two of these points lie together in any connected subset of R of diameter less than ϵ. Let C be a circle having P as center and of diameter $< \epsilon/2$ and such that the exterior of C contains a point Q of R. There exists an integer k and a number e such that for each integer $i > k$, P_i lies within C and is at a distance greater than e from C. Since R is arcwise connected, then for each $i > k$, R contains an arc $P_i Q$. For each $i > k$, on $P_i Q$, in the order from P_i to Q, let X_i denote the first point belonging to C. Then since no two points of the sequence $P_1, P_2, \cdots$ can be joined by any connected subset of R of diameter less than ϵ, it follows that no two of the arcs $P_{k+1} X_{k+1}$, $P_{k+2} X_{k+2}, \cdots$ can have a common point. But each of these arcs is a subset of R and is of diameter greater than e. This is contradictionary to the above mentioned theorem of H. M. Gehman's. Thus the supposition that R does not have property S leads to a contradiction, and our theorem is proved.

THEOREM 15. *Every point of the boundary of an arcwise connected subset of R of a continuous curve M every subcontinuum of which is a continuous curve is regularly accessible from R.*

Proof. It follows by Theorem 14 that R has property S. Then since by a former theorem of the author's,† R is arcwise connected im kleinen,

---

* Cf., for example, the proof of Theorem 1 of my paper just referred to.

† G. T. Whyburn, Proc. Nat. Acad. Sci., vol. 12 (1926), pp. 761–767, Theorem 5.

therefore, by Theorem 13, every point of the boundary of R is regularly accessible from R.

THEOREM 16. *If H denotes the sum of the boundaries of all the complementary domains of a continuous curve M every subcontinuum of which is a continuous curve then H is arcwise connected.*

Proof. Let A be any point of H and N_0 the set of all those points, including A, which can be joined to A by an arc which is a subset of H. Then N_0 exists and contains at least the boundaries of all the complementary domains of M whose boundary contains A. Let $N = N_0 + N_0'$. Then no point of $N - N_0$ can belong to H. For suppose some point P of $N - N_0$ belongs to H. Then since N_0 is arcwise connected and $P \subset N_0'$, then by Theorem 15, P is accessible from N_0, and hence $N_0 + P$ is arcwise connected. And since $P \subset H$ therefore $P \subset N_0$ and not of $N - N_0$, which contradicts our supposition. Thus no point of $N - N_0$ can belong to H.

I shall show that N is identical with M. Suppose it is not. Then some complementary domain R of N contains a subset F of M. Let B denote the boundary of R. Let K denote the continuous curve (possibly unbounded) obtained by adding to $F + N$ all the complementary domains of N except R. Then no point of B belongs to the boundary of any complementary domain of K. For suppose some point P of B does belong to the boundary I of some complementary domain D of K. Since K contains M, clearly D is a complementary domain of M, and furthermore, D is a subset of R. But since R is not a subset of D, it contains at least one point of I. But now I contains the point P of N, and since $H \cdot (N - N_0) = 0$, therefore $P \subset N_0$, and hence $I \subset N_0$. This is impossible, since I contains a point of R, a complementary domain of N. Thus it follows that no point of B can belong to the boundary of any complementary domain of K.

Let C be a circle which, together with its interior, lies wholly in R. Since no point of B belongs to the boundary of any complementary domain of K, then by a theorem of the author's,* for each point P of B there exists a simple closed curve J_{1p} which encloses P and belongs to K but which neither contains nor encloses any point of C. By the Borel theorem there exists a finite number of the simple closed curves $\{J_{1p}\}$ such that every point of B is within at least one of this finite collection of curves. Let Q_1 denote the sum of this finite collection of curves, and let B_1 denote the boundary of that complementary G_1 of the continuum Q_1 which contains C. Then B_1 is a continuous curve. The curve B_1 contains no point of B. For if P is any point of B, then some curve of the collection $\{J_{1p}\}$ which belongs to Q_1 must separate P from C, and hence P is not on the boun-

* Trans. Amer. Math. Soc., vol. 29 (1927), p. 369–400, Theorem 7.

dary of the complementary domain of Q_1 which contains C. Therefore B lies in a complementary domain of B_1 different from G_1. Now since B_1 separates C, a point set belonging to R, from B, the boundary of R, then B_1 must contain a point of R; and since B_1 contains no point of B, therefore B_1 is a subset of R. And as B_1 is a subset of K, and K has in common with R only the point set F, a subset of M, it follows that B_1 is a subset of M.

Now using the last quoted theorem of the author's, it follows that for each point P of B there exists a simple closed curve J_{2p} which belongs to K and encloses P but which neither contains nor encloses any point of B_1. Just as above it follows that there exists a finite collection of these curves such that every point of B is within at least one of them. Let Q_2 denote the sum of the simple closed curves of this finite collection, and let B_2 denote the boundary of the complementary domain of Q_2 which contains B_1. Then just as above in the case of B_1 and C, it is shown that B_2 is a continuum which belongs to M and to R and which separates B_1 and B, i. e., B_1 and B lie in different complementary domains of B_2. Let B_3 be a set selected with respect to B_2 and B just as B_2 was selected with respect to B_1 and B, and so on. This process may be continued indefinitely, and thus we obtain an infinite sequence of mutually exclusive continua B_1, B_2, B_3, $\cdots$ each of which belongs to M and separates C and B. If d is a positive number $< 1/2\,\delta(C)$ and also $< 1/2\,\delta(B)$, then for every positive integer n, $\delta(B_n) > d$. But since every subcontinuum of M is a continuous curve, this is contradictory to the above mentioned theorem of H. M. Gehman's. Thus the supposition that N is not identical with M leads to a contradiction.

Now since N is identical with M, and since, as shown previously, $H \cdot (N - N_0) = 0$, therefore $H \equiv N_0$. But N_0 is, by definition, arcwise connected. Hence H is arcwise connected, and our theorem is proved.

THEOREM 17. *Using the same notation as in Theorem* 16, *if P is any point of $M - H$ and $\varepsilon > 0$, then $M - H$ is the sum of two mutually separated point sets M_1 and M_2, where M_1 contains P and $\delta(M_1) < \varepsilon$.*

Proof. Suppose, on the contrary that there exists an $\varepsilon > 0$ such that $M - H$ is not the sum of any two such sets M_1 and M_2. It is clear that ε may be so chosen that $\delta(M - H) > \varepsilon$. Then since $\delta(H) \geqq \delta(M - H) > \varepsilon$, and, by Theorem 16, H is arcwise connected, H contains an arc t_1 of diameter $> \varepsilon/3$ which does not contain P. Let K_1 denote the component (i. e., the maximal connected subset) of $M - t_1$ containing P, and H_1 the set $K_1 \cdot H$. Then since $K_1 - H_1$ and $(M - H) \cdot (M - K_1)$ are separated, it follows from our supposition that $\delta(H_1) \geqq \varepsilon$. By Theorems 10 and 8, H_1 is connected; and since H_1 is open relative to H and H is* arcwise

* This follows from the fact that H is arcwise connected and a theorem of the author's, Cf. Proc. Nat. Acad. Sci., vol. 12 (1926), pp. 761–767, Theorem 5.

connected im kleinen, then by Theorem 12, H_1 is arcwise connected. Hence H_1 contains an arc t_2 of diameter $>\varepsilon/3$ which does not contain P. Let K_2 be the component of $M-(t_1+t_2)$ containing P, and let $H_2 = K_2 \cdot H$. Then, just as above in the case of H_1, it follows that $\delta(H_2) \geqq \varepsilon$ and that H_2 contains an arc t_3 of diameter $>\varepsilon/3$ which does not contain P. Let K_3 be the component of $M-(t_1+t_2+t_3)$ containing P, and so on. This process may be continued indefinitely, giving an infinite sequence $t_1, t_2, t_3, \cdots$ of arcs in M each of diameter $>\varepsilon/3$. But since every subcontinuum of M is a continuous curve, this is contradictory to H. M. Gehman's theorem quoted above. Thus the supposition that Theorem 17 is not true leads to a contradiction.

COROLLARY. *The set T of all the points of a continuous curve M every subcontinuum of which is a continuous curve which belong to the boundary of no complementary domain of M is totally disconnected.*

THEOREM 18. *Using the same notation as in Theorem* 16, *every point P of $M-H$ is locally isolated by simple closed curves which belong to H.*

Proof. Let C_1 be a circle with center a P and diameter $d/2$. Since P belongs to the boundary of no complementary domain of M, there exists a circle C_2 concentric with and within C_1 and such that no complementary domain of M which has a point on or within C_1 can have a point or a limit point on or within C_2. Let r denote the radius of C_2. By Theorem 17, $M-H$ is the sum of two mutually separated point sets M_1 and M_2, where $M_1 \supset P$ and $\delta(M_1) < r/2$. Let K denote the set of points obtained by adding to $M-H$ the circle C_2 plus its exterior. Then if K_0 denotes the set of points $K-M_1$, it is easy to see that M_1 and K_0 are mutually separated. Then by Knaster and Kuratowski's theorem,* there exists a continuum N with contains no point whatever of K but which separates P from some point Q which is without C_1. Then N lies wholly within C_2. Let T denote the sum of the boundaries of all those complementary domains of M whose boundary contains at least one point of N. Then $T \subset H$, and by an argument similar to, but simpler than that used in the first part of the proof of Theorem 5, it is shown that T is a continuum. Clearly T is within C_1 and hence is of diameter less than d. And since N separates P from Q, it is easily seen that T also separates P from Q. But T, being a subcontinuum of M, is a continuous curve; and by a theorem of R. L. Moore's† it follows that T contains a simple closed curve J which separates P from Q. Since Q is in the unbounded complementary domain of T, then J must enclose P. The truth of Theorem 19 is therefore established.

* Loc. cit.

† Math. Zeits., vol. 15 (1922), pp. 254–260, Theorem 5.

THE UNIVERSITY OF TEXAS.

FERMAT'S NOTE XLV.*

By C. M. Walsh.

By Note XLV is meant the forty-fifth of Fermat's notes written in his copy of Bachet's edition of Diophantus. With this will be compared a passage covering the same subject in Letter CI, so numbered in the Tannery-Henry edition of the *Œuvres de Fermat.*† The Letter was adressed to Carcavi and dated 1659. The dates of the various notes on Diophantus are not known; but it is probable that this Note was written sometime between 1636 and 1641, or at least, in round numbers, twenty years before the Letter.‡ This fact is important; for the Note and the Letter do not agree—at least in appearance.

The common subject of the Note and of the passage in the Letter is the new method of proof by "infinite or indefinite descent", as applied to the proposition that the area of a right-angled triangle in (whole) numbers cannot be a (perfect) square—or in other words, that *in integral solutions of* $x^2+y^2=z^2$, $\frac{1}{2}xy$ *cannot be a square.*

The Note may be literally translated as follows, with a few explanatory interpolations. Numbers are inserted to mark the successive statements.

"If the area of a [right-angled] triangle were a square, (1) two biquadrates would be given whose difference would be a square; whence it follows that (2) two squares would be given such that both their sum and their difference would be squares. Therefore (3) a number equal to a square [i.e. a square number] is given that is composed of a square and the double of a square, with the condition that the squares composing it [taken singly] make a square. But (4) if a square number is composed of a square and the double of another square, its side [= root] is similarly composed of a square and the double of a square, as we can very easily demonstrate. Hence it will be concluded that (5) that side is the sum of the sides about the right angle in a [or the] right-angled triangle, and that one of the squares composing the sum forms the base and the double square is equal to the perpendicular [i. e., the two catheti of this triangle are themselves, the one a single square, the other the double of a square].

* Received November 17, 1927.

† The note is in vol. I, p, 340; the passage in the Letter in vol. II, pp. 431–32.

‡ It was the opinion of P. Tannery (*Mémoires scientifiques*, vol. 6, pp. 15, 17) that Fermat's greatest discoveries, many of which were aided by the method in question, were made between these dates. In fact, for this Note the earlier date is the more likely, for a letter written in 1636 shows that he knew the proposition at that time, *Œuvres*, II, 65.

"(6) That right-angled triangle, therefore, will be made up of two squares whose sum and difference will be squares. But (7) these two squares will be proved to be smaller than the original squares first obtained [in (2)], whose sum as well as difference is a square. Thus, if two squares are given whose sum and difference are squares, there will be given, in whole numbers, a sum of two squares of the same nature [i. e., such that their difference also is a square], smaller than the preceding [sum].

"(8) By the same reasoning there will be given another [sum] smaller than this one, found in the [same] way [as] the one preceding it [was found], and always *in infinitum* smaller whole numbers will be found presenting the same [property]; (9) which is impossible, because when any whole number is given, an infinity of whole numbers cannot be given smaller than it.

"Insertion of the full and detailed demonstration [of the series of ever smaller sums] is prevented by the narrowness of the margin."

The operation described in this Note consists in showing that the hypothesis of $\frac{1}{2}xy$ in $x^2+y^2=z^2$ being a square involves the co-existence of two squares such that both their sum and their difference are squares, and this is impossible because these squares would, by the same reasoning by which they themselves were obtained, involve the co-existence of other squares having this property whose sum would be smaller still, and these would, by the same reasoning repeated, involve the co-existence of still others whose sum would be smaller than theirs, and so on indefinitely. Thus the argument by infinite descent, which is a *reductio ad adsurdum*, is applied to this involved proposition. This being proved impossible as involving an absurdity, the main proposition, which involves it, is proved impossible.

But many years later, in the Letter, Fermat described his proof by descent of this proposition as applied directly to the proposition itself. His words are:

"The proof is made by *ἀπαγωγὴν εἰς ἀδύνατον* in this manner.

"If there were any right-angled triangle [or equation $x^2+y^2=z^2$] in whole numbers that had its area [or $\frac{1}{2}xy$] equal to a square, there would be another [right-angled] triangle [or similar equation] smaller than it having the same property. If there were a second, smaller than the first, having the same property, there would be, by the same reasoning, a third smaller than the second, having the same property; and then a fourth, a fifth, etc., to infinity in descending. Now, it is a fact that when a number is given, there is not an infinity, in descending, of numbers smaller than it (I am speaking always of whole numbers). Whence we conclude that it is impossible that there should be any right-angled triangle whose area is a

square." He adds that the proposition is applicable also to right-angled triangles with sides expressed in fractions,* because these can be turned into triangles with sides expressed in whole numbers; and that he does not give the reason by which he inferred that "if there were a right-angled triangle of this nature, there would be another of the same nature smaller than the first [and so on], because the account of it would be too long† and because therein resides the whole mystery of my method".

The difference between the two descriptions is striking, though it should not be exaggerated.‡ In the Note it is said that the sum of two squares (whose difference also is a square) is a square, and this sum leads down to other similar sums. But these sums of sqares that are squares are nothing more nor less than right-angled triangles. Thus there is no real opposition between the two accounts. It is not necessary to suppose that Fermat later discovered a new and simpler demonstration, which he described in the Letter. He may have had in mind precisely the same demonstration in 1659 that he had in 1636, only in 1659 in describing it, he abbreviated, and left out reference to the accompanying differences of the two squares which also are squares, and spoke of the sums as "triangles", as he also omitted the other details, now writing for others, whereas in this Note he had written only for himself, and had inserted all the various stages of his argument to aid his memory. It is, of course, possible that he may have later discovered a shorter and simpler method of conducting the proof of this proposition by descent. But there is nothing in all his extant writings requiring this supposition. And we ourselves, if we wish to reconstruct Fermat's argument, ought to follow the longer and more detailed account which has been given us by the publication after his death of his own private memorandum. We should follow the Note through all its winding to the end. Now, the seventh and eighth statements are that not only the right-angled triangles, or the sums of two squares equal to squares, are involved in a descending series, but that these sums are such that also the differences of the two squares in all the descending pairs are squares; whereas the Letter refers only to the sums of two squares being squares (right-angled triangles) in a descending line. A demonstration involving only the latter requirement is not the demonstration fully described

* Cf. *Œuvres*, II, 376.

† This reason is out of place in a letter—really a disquisition sent in a letter—where he had at command all the space he needed. Obviously the next is the real reason for withholding a full exposition of his method.

‡ Sir Thomas L. Heath in the second edition of his *Diophantus of Alexandria*, 1910, p. 295, notes a slight divergence, and passes on. Most of the commentators do not even pause to note any difference.

by Fermat. Yet most of the commentators upon the Note have been content with demonstrations that require only the ever descending sums or triangles, in agreement with the meagre account in the Letter.

Not till nearly a century and a half after Fermat's death did the commentators begin, and Legendre was the first. But one of Fermat's correspondents, Frenicle, whom Fermat greatly respected,* used Fermat's method of descent to prove the proposition in question, and while his demonstration carried out the Letter, it did not follow the Note. Yet on this the more recent commentators upon the Note have built. A commentator upon the Letter, G. Wertheim, has inferred that because in a passage at its close†, Fermat spoke of having "sent some demonstrations by descent to Carcavi and Frenicle", therefore the demonstration given by Frenicle was the one (or one of those) sent him, and consequently that it may be taken for Fermat's own demonstration.‡ It is overlooked that a little earlier in the same Letter Fermat mentions a proposition of which he had sent the proof (by descent) to Frenicle—and it was another proposition; and Fermat adds that he sent it because Frenicle had admitted his inability to find it.§ There is no reason, however, to believe that the proof of our proposition was beyond Frenicle's powers; and it does Frenicle an injustice to suppose he would insert in his own work a demonstration obtained from Fermat without acknowledgement.‖ Frenicle could never have seen the Note till it was published in 1670, five years after Fermat's death, when Frenicle's interest in the subject was over. The only other relevant fact we possess is that in 1640 Fermat had submitted to Frenicle the problem of finding a square $\frac{1}{2}xy$ in $x^2+y^2=z^2$, in order to see whether he would perceive its impossibility.¶ But it is next to certain that the Letter (of 1659)

* *Œuvres*, II, 187, 203, 248, 376, 437, with some reserve 252, 344–5.

† *Œuvres*, II. 436.

‡ Wertheim published this opinion in 1899 in the Zeitschrift für Mathematik und Physik, vol. XLIV, Hist. u. Lit. Abtheilung, pp. 4–7. The title of his paper is *Ein von Fermat herrührender Beweis.* He was very positive in his assertion: "Es unterliegt daher wohl keinem Zweifel, daß wir hier mit einem Beweise Fermat's zu thun haben", p. 4. His opinion has been adopted by Heath, *op. cit.*, p. 295. Professor L. E. Dickson maintains the same, *History of the Theory of Numbers*, vol. 2, p. 617.

§ *Œuvres*, II, 434. Frenicle made this admission in a pamphlet (now lost) printed in 1657. Fermat sent him the proof in 1658, as appears from II, 441. In III, 496, Wallis refers to the solution of still another problem which, because of Frenicle's inability, Fermat had sent him.

‖ It was published posthumously, 1676, in Frenicle's *Traité des Triangles en Nombres*, Prop. XXXIX.

¶ *Œuvres*, II, 195. He had tried to test Sainte-Croix in the same way in 1636, p. 65, and later, in 1658, repeated the experiment with Wallis, pp. 346, 376. In 1643 Frenicle suspected that Fermat was again testing him with other impossible problems, and resented it, pp. 260, 262.

was communicated by Carcavi to Frenicle; and the probability is that Frenicle then set himself to use the method by descent and succeeded in working out his demonstration. It is, of course, quite possible that it resembled Fermat's own demonstration, if, as is possible, Fermat then had a new one simpler than the one he described in the Note. But, to repeat, we have no reason to believe this. That Frenicle's demonstration was his own, is a conjecture on its face as good as Wertheim's, while internal evidence will show it to be better.

As the beginning and the ending of Frenicle's demonstration (which we may believe to have been made entirely without reference to the Note) lie at the bottom of the commentators' efforts to unravel Fermat's procedure as described, not merely in the Letter, but in the Note, it deserves to be briefly reviewed.

Frenicle knew that it was sufficient to prove the proposition of primitive solutions only, and that all primitive solutions of

(1) $$x^2+y^2=z^2$$

are given by the formulae

$$x=p^2-q^2, \qquad y=2pq, \qquad z=p^2+q^2,$$

where the generating numbers (as he called them) p and q are any integers prime to each other and odd and even.* Now, if $\frac{1}{2}xy=\square$, then $pq(p^2-q^2)=\square$, and as pq and p^2-q^2 also are prime to each other, pq and p^2-q^2 must both be squares and $2pq$ the double of a square. Let $x(=p^2-q^2)=a^2$. Thus we have a second equation

(2) $$a^2+q^2=p^2,$$

where a and p are odd, and q even. Furthermore, since pq is a square and p and q are prime to each other, both p and q are squares. The equation (2) is completely solved, as to primitive solutions, by

$$a=p_1^2-q_1^2, \qquad q=2p_1q_1, \qquad p=p_1^2+q_1^2,$$

where again the generators p_1 and q_1 are integers prime to each other and odd and even.† Since $q=2p_1q_1=\square$, one of the numbers p_1 and q_1 is an odd square and the other the double of a square. As p is a square, let $p=u^2$. Thus we have a third equation

* It is possible to interchange the formulae of x and y, making x the even and y the odd number. But nothing is gained thereby. Frenicle knew that in primitive solutions of the equation x and y must be odd and even (z always being odd), and so he could choose which he pleased for the odd and which for the even.

† The formulae of a and p here are not interchangeable.

(3) $$p_1^2 + q_1^2 = u^2,$$

in which the hypothesis is repeated, $\frac{1}{2} p_1 q_1$ being a square.* In this equation the numbers are all smaller than the corresponding ones in (2); for p_1 is $< (p_1 + q_1)(p_1 - q_1) = a$, $q_1 < 2 p_1 q_1 = q$, and $u < u^2 = p$. And the numbers in (2) are smaller than the numbers in (1); for a is $< a^2 = x$, $q < 2pq = y$, and $p < p^2 + q^2 = z$. Therefore the numbers in (3) are smaller than those in (1).

Thus from the original equation (1), on the hypothesis that $\frac{1}{2} xy = \square$, has necessarily been derived another equation (3) with smaller corresponding numbers, and with the half of the product of the first two terms a square, which renews the hypothesis concerning the original equation. And by the same reasoning, Frenicle now says, we could obtain another equation smaller than (3) and another still smaller, having the same property as (3) and as (1), and so on, forever descending; which is impossible in integers.†

Here, though Frenicle's argument really covers several of Fermat's statements in the Note, he ignores them as side-issues. Especially he has no use for the third, fourth and fifth, which played so great a rôle, apparently, in Fermat's demonstration. And his ending changes the last three statements into something very different: not the sums of the squares whose sums and differences are squares are ever smaller and smaller (for it cannot be proved that $p_1^2 - q_1^2 = \square$), but the original equation (or triangle) itself is shown to be followed by an endless succession only of similar smaller equations (or triangles). Thus it is not the proof as described in the Note, even though it should be, or be like, Fermat's own proof, if Fermat later had one that was sufficiently described in the Letter. Frenicle was not elaborating the Note: he was following out the indications given in the Letter, and he succeeded in doing so without requiring much that was, apparently, required in the demonstration described in the Note. We may,

---

* In this equation if we put $p_1 = m^2$ and $q_1 = 2n^2$, it becomes $(m^2)^2 + (2n^2)^2 = u^2$, which illustrates Fermat's fifth statement. (If further we put $2n = k$, the equation becomes $(m^2)^2 + (\frac{1}{2} k^2)^2 = u^2$, a form which will appear later.) Frenicle did not go out of his way to note these possible forms, which served no purpose in his argument. Fermat is famous for the conciseness of his expositions. If the several statements in his Note were not necessary for his argument, why did he insert them—especially on the margin of a page, where conciseness was especially needed?

† For Frenicle's argument it is not necessary that equation (1) should have the property possessed by equation (3) that the first term is a square and the second the double of a square, but as Fermat laid emphasis upon this property, it would seem to be required in his argument. Nevertheless, even in Frenicle's exposition, this property is possessed by equation (1); for it was shown that both $p^2 - q^2$ and pq must be squares, which means that x must be a square and y the double of a square.

therefore confidently assert that Frenicle's demonstration was not Fermat's, as fully described in the Note. And to suppose it was like a later one attained by Fermat, we have no reason at all.

Coming now to the commentators, we find that Legendre, while trying to explain the Note, pretended to do no more than to carry out its various stages with sufficient exactness ("assez strictement")*. He began as Frenicle did, but went into greater detail. Finding that on the supposition of $\frac{1}{2}xy = \square$ in $x^2+y^2 = z^2$ the generators of the primitive solutions of this equation must be squares, as also the difference of their squares, he put $p = u^2$ and $q = v^2$,† but merely remarked that then $p^2 - q^2 = \square$ becomes $u^4 - v^4 = \square$, which represents Fermat's first statement. And now as $u^4 - v^4$ has the two factors u^2+v^2 and u^2-v^2, and as u and v are prime to each other (as squares of the relatively prime p and q) and (like p and q) are odd and even, therefore the factors u^2+v^2 and u^2-v^2 are prime to each other; and as their product is a square, they must each be a square; which was Fermat's second statement. Putting

$$(1) \qquad u^2+v^2 = r^2 \quad \text{and} \quad u^2-v^2 = s^2,$$

he now got, by subtracting the second from the first, and transposing the terms in the second,

$$s^2+2v^2 = r^2 \quad \text{and} \quad s^2+v^2 = u^2,$$

which two constitute Fermat's third statement. As for Fermat's fourth statement, Legendre referred back to an earlier part of his work where he himself had proved it.‡ This leads him to re-discover an algebraic solution of the equation $s^2+2v^2 = r^2$ which Euler had previously given and which is known to be complete as to primitive numerical solutions. This is contained in the formulae

$$s = m^2-2n^2, \qquad v = 2mn, \qquad r = m^2+2n^2,$$

which Legendre tacitly supposes Fermat to have known.§ On substituting these values of s and v in $s^2+v^2 = u^2$, he obtained $m^4+4n^4 = u^2$, which should be possible if $\frac{1}{2}xy = \square$. This is the same as

* *Essai sur la Théorie des Nombres*, 1st ed., 1798, §§ 334–5, see the foot-note on p. 402.

† The letters used in this paper are not necessarily those of the authors reviewed. They are adopted so as to be uniform, as far as possible, in all the various elaborations.

‡ § 141. He did not assert that his proof was the same as Fermat's.

§ Fermat could not have obtained this solution even in the same way as Euler did. Professor Dickson will tell us, in a foot-note further on, how Fermat could have reached $r = m^2+2n^2$ and $s = m^2-2n^2$, after which the value of v is apparent.

(2) $$(m^2)^2+(2n^2)^2 = u^2,$$

which is Fermat's fifth statement. Here $\frac{1}{2}m^2(2n^2) = (mn)^2$, and so it repeats the hypothesis concerning the original equation, the area of this triangle likewise being a square.

Legendre now skips Fermat's sixth statement, and does not notice the divergence of his next statement from Fermat's seventh and eighth. His own next statement is that in the last equation the numbers m^2 and $2n^2$ and u are smaller—and he emphasizes that they are much smaller*—than the numbers in the original $x^2+y^2=z^2$. Then, lastly, like Frenicle, he argues that, having shown that, if a solution of $x^2+y^2 = z^2$ exists with $\frac{1}{2}xy$ a square, there must also exist another solution with $\frac{1}{2}xy = \square$, but smaller than the preceding, so, by the same reasoning as that by which this was proved, he can prove that there must be a third still smaller solution in which $\frac{1}{2}xy$ is a square, and so on *in infinitum*; which is impossible in a descending line in integers.

Here Legendre avoids any mention, as did Fermat himself, of an intervening smaller triangle (or equation) between the original $x^2+y^2=z^2$ and the smaller one which resembles it in having its area a square. Legendre's demonstration of the proposition thus more closely follows Fermat's description, not only in the Note, but also in the Letter, than did Frenicle's. As for the Note, Legendre omitted only one of Fermat's statements, the sixth, though he had no further use for the second, which Fermat used so frequently; and consequently, from there on, his ending is different from the ending of Fermat's demonstration as described in the Note; for Legendre did not, and could not, show that $(m^2)^2-(2n^2)^2 = \square$.

Legendre added that the same demonstration proves the impossibility in integers (except certain trivial cases) of the equations $u^4-v^4 = \square$ and $m^4+4n^4 = \square$ in general (because if the former were possible, on using the particular u^2 and v^2, the difference of whose squares is a square, for the generators of a solution of $x^2+y^2=z^2$, we should have a solution in which $\frac{1}{2}xy[= u^2v^2(u^4-v^4)]$ is a square;† and if the second were possible, we should directly have an equation, $(m^2)^2+(2n^2)^2 = \square$, in which $\frac{1}{2}m^2 2n^2$ is a square). Also, he said, it follows that the equation $x^4+y^4 = 2z^2$ is impossible, because (on squaring both sides and sub-

* He shows that $(mn)^2 < \sqrt[3]{\frac{1}{2}xy/128}$.

† According to Dickson, *op. cit.*, II, 618–19, A. Ginocchi condemned this proof as incomplete, because u and v are odd and even, whereas in the equation in question they might both be odd. But if the equation with u and v both odd were possible, by using their squares for generators we should get derivitive solutions of $x^2+y^2=z^2$ with $\frac{1}{2}xy = \square$; which is equally impossible.

tracting $4x^4y^4$ from both sides and dividing by 4) we should have $\{\frac{1}{2}(x^4-y^4)\}^2 = z^4-(xy)^4$, which has just been proved impossible. The first of these immediately shows the impossibility of $x^4-y^4 = (z^2)^2$ and consequently of $x^4+y^4 = z^4$ in general. This consequence is so plain that if Fermat had worked out his demonstration in this way, we should wonder why he did not append it to his Note. On the contrary, in Note XXXI* he asserts that $x^4+y^4 = z^4$ is impossible in integers "as our [my] method of demonstrating can undoubtedly disclose". "Our method" here seems to be the method by descent, just as, in the Letter, he explicitly tells us that he used this method to prove the impossibility of $x^3+y^3 = z^3$, but in a way entirely different from that used in proving the impossibility of $\frac{1}{2}xy = \square$ in $x^2+y^2 = z^2$.† It is likely that he proved the impossibility of both $x^4+y^4 = z^4$ and $x^3+y^3 = z^3$ before he was able to reach, and to prove, his "great theorem" (enunciated in Note II) about the impossibility of $x^n+y^n = z^n$, $n>2$.‡ Now, we know that Euler later proved both these particular impossibilities by the method of descent, and the latter in a manner very different from the former. Moreover, Fermat in his extant writings shows no knowledge of the impossibility of the other equations signalized by Legendre. The inference seems to be that Fermat did not use Legendre's method of proving the proposition.

Another attempt to interpret the Note, without acknowledgement of the preceding attempt, but following it mostly, and yet with a few important divergences, was made by H. G. Zeuthen.§ As Zeuthen's interpretation has been adopted word for word by Sir Thomas L. Heath,‖ and has been followed, again with slight divergence, by Professor L. E. Dickson, it may be well to review it in full.

Like Frenicle and Legendre, Zeuthen used the formulae giving all primitive solutions of $x^2+y^2 = z^2$, in which not only p and q are prime to each other, but, because they are odd and even, also $p+q$ and $p-q$ are prime to them and to each other. The hypothesis $\frac{1}{2}xy = \square$ requires $pq(p^2-q^2) = pq(p+q)(p-q)$ to be $= \square$, where all the four relatively prime factors must be squares. Let

(1) $p = u^2$, $q = v^2$, $p+q = u^2+v^2 = r^2$, and $p-q = u^2-v^2 = s^2$.

---

* *Œuvres*, I, 327.

† *Œuvres*, II, 453.

‡ It is probable, too, that Note XXXI was written *after* Note XLV; for if he had known the impossibility of $x^4+y^4 = z^4$ when he tried to prove the impossibility of $\frac{1}{2}xy = \square$ in $x^2+y^2 = z^2$, why did he not stop at his very first statement?

§ *Geschichte der Mathematik im XVI. und XVII. Jahrhundert*, 1903, pp. 163–4.

‖ *Op. cit.*, 294–5.

Here the system $u^2+v^2=r^2$ and $u^2-v^2=s^2$ is Fermat's second statement. And on multiplying these equations by each other, member for member, we have

(2) $$u^4-v^4=(rs)^2,$$

which is Fermat's first statement, although here the order of the statements is reversed unnecessarily; for Zeuthen might equally well have followed Legendre's order. By subtracting $u^2-v^2=s^2$ from $u^2+v^2=r^2$ we get

(3) $$2v^2=r^2-s^2=(r+s)(r-s),$$

and so we have Fermat's third statement ($s^2+2v^2=r^2$, where, from (1), $s^2+v^2=u^2$).

Now, $r+s$ and $r-s$ are both even numbers, because r^2 and s^2 must both be odd; but $r+s$ and $r-s$ can have no other common factor except 2, since u^2 and v^2 (like p and q) are prime to each other, and consequently also r and s. Zeuthen now says it follows that

$$r+s=\begin{cases}2m^2\\k^2,\end{cases}\quad\text{and}\quad r-s=\begin{cases}k^2\\2m^2,\end{cases}$$

where k is an even number; and also

$$u^2=\frac{1}{2}(r^2+s^2)=(m^2)^2+\left(\frac{1}{2}k^2\right)^2.$$

He does not indicate how he obtained these results. We may supply the intervening links as follows. The generator q has been taken for the even one, and therefore v^2 is even, and we may put $v=2l$, and therefore (3) becomes $8l^2=(r+s)(r-s)$, or

$$(2l)^2=(r+s)\left\{\frac{1}{2}(r-s)\right\}\quad\text{or}\quad=\left\{\frac{1}{2}(r+s)\right\}(r-s),$$

where $r+s$ and $\frac{1}{2}(r-s)$ are prime to each other, as also $\frac{1}{2}(r+s)$ and $r-s$. Hence we must have

(4) $$r+s=2m^2\text{ or }k^2\quad\text{and}\quad r-s=k^2\text{ or }2m^2,$$

where k is even. It is sufficient to adopt one of these alternatives, and Zeuthen chose the first. From (1) we get $u^2+s^2=v^2$ and from (3) $v^2=\frac{1}{2}(r^2-s^2)$, whence $u^2=\frac{1}{2}(r^2+s^2)$; and now, using the first values in (4), we have $r^2+2rs+s^2=4m^4$ and $r^2-2rs+s^2=k^4$, whence by addition and halving, $r^2+s^2=2m^4+\frac{1}{2}k^4$, whence again

$$u^2 = \frac{1}{2}(r^2+s^2) = (m^2)+\left(\frac{1}{2}k^2\right)^2, \tag{5}$$

the final result obtained by Zeuthen. Thus the whole numbers m^2 and $\frac{1}{2}k^2$ are numbers the sum of whose squares is equal to another square, as in the original equation, and half their product is the square $(\frac{1}{2}mk)^2$. That is, we have over again the form $x^2+y^2=z^2$ in which $\frac{1}{2}xy = \square$.

Zeuthen offers his exposition as carrying out Fermat's fourth statement, because from (3) we have $s^2+2v^2=r^2$, and $r = m^2+2(\frac{1}{2}k)^2$ is obtained from the two equations in (4) by using the first alternative, so that r has the same form as r^2. But Fermat did not say that as a consequence of $\frac{1}{2}xy$ being $= \square$ in $x^2+y^2=z^2$ he could obtain an instance of $s^2+2v^2=r^2$ in which r has the same form as r^2: he made the general statement that in all equations $s^2+2v^2=r^2$, r has the same form as r^2,—and asserted that he could easily prove it. This general statement is not proved in the above procedure, and no use is made of Fermat's knowledge of this general proposition. Thus Zeuthen does not render Fermat's statement, or follow Fermat's procedure.

Zeuthen's next statement, however, is correct. For k must be even, and if we put $k=2n$, the equation in (5) becomes $m^4+4n^4=u^2$ or

$$(m^2)^2+(2n^2)^2 = u^2, \tag{6}$$

in which one side of the right-angled triangle is a square and the other the double of a square, as in Fermat's fifth statement.* Zeuthen has thus

* It may here be noted that Professor Dickson, *op. cit.*, pp. 615–16, took advantage of Zeuthen's jump before reaching (4) and (5) to obtain (6) directly without using (4) and (5), (at the same time proving generally Fermat's fourth statement differently from the way Legendre and Euler, and Bachmann, had proved it, and in a way he thought more likely to have been Fermat's own,—so in an earlier paper in the Bulletin of the American Mathematical Society, XVII, 1911, p. 532, to which he refers). Having followed Zeuthen down to (3), which he put in the form $s^2+2v^2=r^2$, he set $\frac{1}{2}(r+s)=e$ and $\frac{1}{2}(r-s)=f$, where e and f are integers and $ef=\frac{1}{2}v^2$. A common factor of e and f would divide r, s, u^2 and v^2, which is impossible, as these are relatively prime. We may take e odd (changing, if necessary, the sign, of s). Therefore in $2ef=v^2$, with e an odd square, say $=m^2$, f must be the double of a square, say $=2n^2$; and so $r=e+f=m^2+2n^2$ (which justifies Fermat's fourth statement). He also has $s=m^2-2n^2$. Hence $u^2(=v^2+s^2)$ $=4m^2n^2(=2ef)+m^4-4m^2n^2+4n^2=(m^2)^2+(2n^2)^2$, which is Zeuthen's (6), where the area of the triangle is $(mn)^2$. The rest of Dickson's treatment is not clear. He says the equation $(m^2)^2+(2n^2)^2=u^2$ is completely solved by $m^2=p_1^2-q_1^2$, $2n^2=2p_1q_1$, $u=p_1^2+q_1^2$, in which again p_1 and q_1, relatively prime, and their product a square, must be squares, so that we can put $p_1=u_1^2$ and $q_1=v_1^2$. Now, from $p_1q_1=n^2$ we get $u_1v_1=n$, a factor of $2mn$ or v. Hence u_1 and v_1 are each smaller than v, and consequently (because of $u^2-v^2=s^2$) smaller than u. Here he ends.

arrived, by an unnecessarily long process, at Legendre's (2), which was latent in Frenicle's (3). Zeuthen now ends, as Legendre and Frenicle did, by pointing out that this triangle is smaller than the original triangle, and can be treated in the same way, leading to smaller triangles without end; which is impossible in integers. And he adds, with Legendre, that the impossibility of $x^4 - y^4 = z^2$ is contained in this proof.

Thus Zeuthen, like Legendre, has no use for Fermat's second statement and skips the sixth; and his ending, like Legendre's and Frenicle's, is unlike Fermat's, for again he cannot prove that $(m^2)^2 - (2n^2)^2 = \square$. He, too, does not thoroughly interpret the Note.*

We may now go back somewhat in time and take notice that shortly after Legendre Peter Barlow, who trod to some extent, but not altogether in Legendre's foot-steps, gave a proof of the main proposition before us, which has some resemblance to Fermat's procedure, although he was making no attempt to interpret Fermat's Note.† Barlow's proof commences with what we may take as a

Lemma: *The system* $x^2 + y^2 = z_1^2$ *and* $x^2 - y^2 = z_2^2$ *is impossible in integers,* $y > 0$. This he proves by Fermat's method of infinite descent, briefly thus. He shows that the generators of these two equations give $p_1^2 - q_1^2 = p_2^2 + q_2^2$ and $p_1 q_1 = p_2 q_2$, where for instance p_2 must divide $p_1 q_1$; wherefore we may put $p_2 = ab$ and $p_1 = ap_3$ and $q_2 = bq_3$; whence $a^2 p_3^2 - b^2 q_3^2 = a^2 b^2 + p_3^2 q_3^2$; whence, again, after a transposition, $(a^2 + q_3^2)/(a^2 - q_3^2) = p_3^2/b^2$; and because of the relative primality of all these terms, at least when the first two are divided by 2, therefore we have either

$$a^2 + q_3^2 = p_3^2 \quad \text{and} \quad a^2 - q_3^2 = b^2,$$

or, from $\frac{1}{2}(a^2 + q_3^2)/\frac{1}{2}(a^2 - q_3^2) = p_3^2/b^2$, after another transposition,

$$p_3^2 + b^2 = a^2 \quad \text{and} \quad p_3 - b^2 = q_3^2;$$

which systems are smaller than the original one; and so on forever descending; which is impossible.‡

---

* Zeuthen called his elaboration the *Vervollständigung* of Fermat's demonstration. I am in possession of the late J. L. E. Dryer's copy of Zeuthen's work. In the margin he has marked *Verbesserung* as a better word.

† *An Elementary Investigation of the Theory of Numbers,* 1811, pp. 109-12, 121.

‡ In 1874 G. Le S. Destournelles gave another proof by descent of this lemma, according to Dickson, *op. cit.,* II., p. 467. The lemma was known to Leonardo of Pisa early in the thirteenth century, and to earlier Arabian mathematicians, probably through their inability to find a "congruent number" (y in $x^2 \pm y = \square$) that was a square. Barlow's is, apparently, the first direct proof of it, though it was proved indirectly through the earlier proof of the impossibility of $x^4 - y^4 = z^2$.

Corollaries: *The system* $x^2+y^2=2z_1^2$ *and* $x^2-y^2=2z_2^2$ *and the system* $x^2+4y^2=z_1^2$ *and* $x^2-4y^2=z_2^2$ *are impossible in integers.*

The first, because then $z_1^2+z_2^2=x^2$ and $z_1^2-z_2^2=y^2$, which is impossible; and the second, because then $z_1^2+z_2^2=2x^2$ and $z_1^2-z_2^2=2(2y)^2$, which is impossible.

Main Proposition: *In* $x^2+y^2=z^2$ *in integers* $\frac{1}{2}xy$ *cannot be a square.* For on putting $\frac{1}{2}xy=w^2$, then $z^2+4w^2=(x+y)^2$ and $z^2-4w^2=(x-y)^2$, which by the second of the Corollaries is impossible.

Barlow could easily have avoided recourse to the Corollary, by putting $2xy$ (instead of $\frac{1}{2}xy$) $=w^2$, which would have given him $z^2+w^2=(x+y)^2$ and $z^2-w^2=(x-y)^2$, contrary to the Lemma itself. Thus Barlow's demonstration virtually consists in showing that the supposition of $\frac{1}{2}xy$ in $x^2+y^2=z^2$ being a square involves the co-existence of two squares whose sum and difference are squares; and this he proves impossible by Fermat's method of infinite descent. Now, we saw above that Fermat's demonstration as described in the Note, proceeds by showing that such a supposition involves the co-existence of two such squares, whose sum and difference are squares; which he then shows involve the co-existence of two smaller squares having the same property, and so on descending without end. But the resemblance is spoilt by Fermat's words "by the same reasoning" occuring where they do in his exposition; for Barlow proved by infinite descent the general proposition that the system of two such squares is impossible, first because they lead to a smaller system, and then, by the same reasoning, to another still smaller, and so on; whereas Fermat proved only that the special squares led to by the supposition are impossible because in accordance with the supposition these lead down to lower similar squares, and these again, by the same reasoning (involving all the stages that were needed for the first descent), lead down to still lower, and so on endlessly. Nowhere in his extant writings does Fermat show knowledge of the impossibility of the system $x^2\pm y^2=\square$ in general. At all events, Barlow had absolutely no need of most of Fermat's statements between the first and the sixth.*

Because of the failure of all attempts by previous commentators to expound Fermat's Note fully and consistently, I venture to offer a new exposition. I begin by adopting the artifice used by Barlow, as amended above. This opening is the simplest one possible, and therefore not unlikely the one entered upon by Fermat.

I imagine, then, that Fermat started out by observing that if in $x^2+y^2=z^2$ $\frac{1}{2}xy$ is a square, also $2xy$ is a square. We may put $2xy=w^2$.

---

* It is to Barlow's demonstration that Dryer's word *Verbesserung* is most appropriate.

On adding and subtracting these equal values to and from the two sides of the equation, we get $(x \pm y)^2 = z^2 \pm w^2$, that is, we have the system

$$(1) \qquad z^2 + w^2 = (x+y)^2 \quad \text{and} \quad z^2 - w^2 = (x-y)^2,$$

and on multiplying these by each other, member for member, we get

$$(2) \qquad z^4 - w^4 = (x^2 - y^2)^2.$$

Here we have Fermat's first two statements, only in reverse order; which may be explained by carelessness on Fermat's part (as was later shown by Zeuthen), or may have another explanation to be noticed presently. In (1) on subtracting the second from the first, and transposing, we get

$$(3) \qquad (x-y)^2 + 2w^2 = (x+y)^2,$$

wherein, as a simple identity, we have

$$(x-y)^2 + w^2 = (x^2 + y^2) = z^2;$$

which two constitute Fermat's third statement.

Now, the equation (3) has been solved by Euler* thus:

$$x - y = a^2 - 2b^2, \qquad w = 2ab, \qquad x + y = a^2 + 2b^2;$$

for any two odd numbers can be represented by $x \pm y$. Euler did not state it, but it is shown by the way these formulae were obtained, that they yield all primitive solutions when a and b are any relatively prime integers and a is odd.† This shows that in all solutions the number $(x+y)$ whose square, in (3), is made up of a square and the double of a square, is itself made up of a square and the double of a square. This, his fourth statement, Fermat asserted he could easily demonstrate. This statement, as already remarked, is perfectly general. It is evident, then, that if he knew this last formula, he knew the others too. We may safely, therefore, especially as we have Legendre's precedent, credit Fermat with knowledge of this solution, prior to Euler.

Here, on adding together this formulae of $x - y$ and $x + y$, and on subtracting the former from the latter and halving the results, we get

$$x = a^2 \quad \text{and} \quad y = 2b^2.‡$$

---

* *Vollständige Anleitung zur Algebra*, II, ii, § 185.

† There is a variant solution, also given by Euler, viz., $x - y = 2a^2 - b^2$, $w = 2ab$, $x + y = 2a^2 + b^2$. These merely interchange a and b, and alter the signs; all which is indifferent.

‡ If we used the alternative formulae mentioned in the preceding foot-note, we should get $x = 2a^2$ and $y = b^2$; which merely means that x is now the even number and y the odd.

In these formulae we have Fermat's fifth statement, that there is a triangle whose perpendicular sides are a square and the double of a square, and the sum of whose sides ($a^2+2b^2 = x+y$) is the sum of the sides of a right-angled triangle ($x^2+y^2 = z^2$). Only, in this elaboration of his procedure, this triangle is the original triangle itself. It proves, however, that in the original triangle on the supposition that $\frac{1}{2}xy$ is a square, x must be a square and y the double of a square.*

Again, $(a^2)^2+(2b^2)^2 = z^2$, taking the place of $x^2+y^2 = z^2$, is completely solved, as to primitive solutions, by

$$a^2 = p^2-q^2, \qquad 2b^2 = 2pq, \qquad z = p^2+q^2,$$

where p and q are integers prime to each other and odd and even. Therefore, because $pq = b^2$, both p and q must be squares, and we may put $p = u^2$ and $q = v^2$, u and v likewise being integers prime to each other and odd and even. Then

$$u^4-v^4 = a^2, \tag{4}$$

where the factors of the first member, u^2+v^2 and u^2-v^2 are prime to each other. Therefore these must both be squares, and so we may put

$$u^2+v^2 = r^2 \quad \text{and} \quad u^2-v^2 = s^2, \tag{5}$$

which is Fermat's sixth statement. And on comparing (5) with (1) we see that here r^2, which is smaller than p^2+q^2 and consequently than z, must be smaller than $(x+y)^2$ which is greater even than z^2. Thus the sum here, u^2+v^2, is smaller than the sum there, z^2+w^2; but *it cannot be proved* that the difference here, u^2-v^2, is smaller than the difference there, z^2-w^2, for there $(x-y)^2$ might be as small as 1; all which is precisely Fermat's seventh statement. For in that statement Fermat was careful to say only that the sum in the second system is smaller than the sum in the first.

This elaboration of Fermat's procedure follows, and carries out, all his statements down to and including the seventh, better than any of the previous ones. It remains to show that it carries out the eighth statement, which none of the previous interpretations have done, by showing how a still smaller system is "by the same reasoning" required to be given, and so on forever. This also it can do.

---

* This could much more easily be proved directly. For in primitive solutions of $x^2+y^2 = z^2$ x and y are prime to each other and odd and even. Therefore, if $\frac{1}{2}xy = \square$, $xy = 2\square$; and if x be the odd number, x must be a square and y the double of a square.

We may begin with (4), which has immediately given us (5), now in the same order as Fermat's first and second statements. Perhaps this explains why Fermat mentioned them in this order; for when he wrote his description, this second start may have been more in his mind. In (5), on subtracting the second equation from the first, and transposing, we get

$$s^2+2v^2 = r^2,$$

where also, from the second in (5),

$$s^2+v^2 = u^2,$$

which constitute Fermat's third statement over again. Carrying out his fourth statement, we know that the first of these is completely solved, as to primitive solutions, by

$$s = m^2-2n^2, \qquad v = 2mn, \qquad r = m^2+2n^2,$$

where m and n are integers prime to each other and m is odd. Substituting these values of s and v in the second, we get

$$(m^2)^2+(2n^2)^2 = u^2,$$

which is Fermat's fifth statement again. This equation is completely solved, as to primitive solutions, by the familiar formulae

$$m^2 = p_1^2-q_1^2, \qquad 2n^2 = 2p_1q_1, \qquad u = p_1^2+q_1^2,$$

where p_1 and q_1 are integers prime to each other and odd and even. Here again, because of $p_1q_1 = n^2$, p_1 and q_1 must both be squares, so that we may put $p_1 = u_1^2$ and $q_1 = v_1^2$, and then

$$u_1^4-v_1^4 = m^2,$$

in which the factors of the first member, $u_1^2+v_1^2$ and $u_1^2-v_1^2$, are prime to each other, and must therefore be squares, so that we may put

$$u_1^2+v_1^2 = r_1^2 \quad \text{and} \quad u_1^2-v_1^2 = s_1^2,$$

which is Fermat's sixth statement once more. And now $u_1^2+v_1^2 = r_1^2$ is obviously smaller than $u^2+v^2 = r^2$ in (5), repeating the seventh statement. Here we are back again where we were at (4) and (5), only on a lower level. And henceforth "the same reasoning" may be repeated according

to the eighth statement, always running through Fermat's several statements in the same order, and leading downward forever, and so justifying the ninth statement, which declares the whole process impossible from the beginning, because it can not continue without an end.

An objection may be raised that the demonstration, as thus expounded, is the most complicated of all, despite the easiness of its beginning. Yet this may be turned in its favor; for discoverers generally pursue at first the line of least resistance, only to find afterward very often, themselves or others, that another course, less obvious, is on the whole simpler. Thus Frenicle, if not Fermat himself, about 1659 found an easier demonstration, still by the method of infinite descent, than the one described twenty years before. In this exposition of that earlier demonstration, moreover, the impossibility of $x^4-y^4=z^2$ in general is not obtrusively apparent, as it was in Legendre's and Zeuthen's elaborations; nor was the impossibility of the system $x^2\pm y^2=\square$ in general shown, which explains why Fermat did not fall upon the method adopted by Barlow. If this exposition be the correct one of the demonstration first employed, and described in the Note, it leaves the question open as to whether Fermat himself later simplified the demonstration in the way set forth in the Letter, or whether this was merely an abbreviated account of the same complex operation. All we can say is that we have no evidence that the Letter necessarily refers to a simpler demonstration, like that employed by Frenicle.

After all, the most astonishing part of the whole matter is that the great method of proof by infinite or indefinite descent, of which Fermat was so proud and from which he expected such great results,* was entirely unnecessary for proving the proposition to which he in the Note applied it, and which was, probably, the first proposition where he, so unnecessarily, found need of it. Fermat challenged Wallis to prove that in $x^2+y^2=z^2$ $\frac{1}{2}xy$ cannot be a square,† and Wallis immediately gave a very short and simple proof of it—only too much so. He first made an original derivation of his own of the usual formulae completely yielding primitive solutions of the equation (which was uncalled for), and then in a couple of lines declared that $pq(p^2-q^2)$ cannot be a square because pq is a "plane number" and pq and p^2-q^2 cannot be "similar plane numbers", and therefore their product cannot be a square.‡ For not proving the latter impossibility he was taken to task by an anonymous writer, most likely Frenicle.§ From

* *Œuvres*, I, 340, II, 313, 431, 436.

† *Œuvres*, II, 376.

‡ In letter no. 44 of his *Commercium Epistolicum*, 1658, (in his *Opera Mathematica*, II, 956; also in Fermat's *Œuvres*, III, 600).

§ In Fermat's *Œuvres*, III, 609.

the allusions to "plane" and "similar plane numbers" it is evident that Wallis was relying on Euclid. Euclid indeed, in the *Elements*, IX, 1 and 2, showed that the product of two different numbers can be a square only if they are "similar plane numbers", which, unless they both be squares, always have a common divisor; and as regards square numbers, such as p^2 and q^2, he proved not only that they always have an integral mean proportional, viz. pq (VIII, 11), but also that if they are prime to each other, the proportion $p^2 : pq = pq : q^2$ is in its lowest terms (VIII, 1 and 2 Cor.), and then, principally by means of VII, 28, he proved, in IX, 15, that in this proportion when in its lowest terms (that is, always when p^2 and q^2 are prime to each other, as they are here) pq and p^2+q^2 are prime to each other. This last proposition, like VII, 28, also, is obviously not enunciated as fully as it might be, since it is provable, by very similar reasoning, that pq and p^2-q^2 are prime to each other. Consequently pq and p^2-q^2 cannot be "similar plane numbers", and their product cannot be a square—unless they both are squares. Then, with $p = p_1^2$ and $q = q_1^2$, $p_1^4-q_1^4$ must be a square, and Wallis should have proved that this is impossible. Because of his not proving this, his proof was defective. But it can be proved—and, too, without the method by infinite descent,—and so his mode of proving the proposition was sound. And shortly after Fermat's death Leibniz tried to prove the same proposition, and did so by showing that the supposition which it denies leads to the necessity of there being three square numbers in arithmetical progression with their difference a square, which he pronounced "absurd".* Here, too, was an involved proposition which needed proof, and the fact that Leibniz failed to prove it may have been the reason why he did not publish his demonstration. But as it can be proved—and again without the method by infinite descent,†—his mode of proving the proposition also was sound.

There are three places in Fermat's own demonstration at which he might have shortened his labors by employing a more direct proof than the method by descent. That he did not do so may perhaps be explained by his avowal that "le voies ordinaires me lassent".‡ For in his account of it he showed, as we have seen, that the supposition of $\frac{1}{2}xy$ in $x^2+y^2=z^2$ being a square leads to three necessary possibilities in integers: (1) of

---

* His paper, written in 1678, was not published till 1863 in Gerhardt's edition of *Leibnizens Mathematische Schriften*, vol. VII, pp. 120–125.

† Fermat himself knew, or suspected, this impossibility, *Œuvres*, II, 63. Who first proved it, it would be interesting to know. It is easy to prove after proving the impossibility of the system $x_2 \pm y^2 = 2\square$. Barlow set it as an exercise for his pupils, *op. cit.*, p. 257 having proved the impossibility referred to. Another proof will be given presently.

‡ *Œuvres*, II, 187.

$x^4 - y^4 = z^2$, (2) of the system $x^2 + y^2 = z_1^2$ and $x^2 - y^2 = z_2^2$, and (3) of the system $x^2 + 2y^2 = z_1^2$ and $x^2 + y^2 = z_2^2$. If, then, he had proved any of these things to be impossible, he would have proved his proposition. We have seen that later Euler and Barlow proved the impossibility of the first and second respectively, and, it may be added, Barlow proved the impossibility also of the third,* but both of them always using Fermat's own method by descent. They can be proved without it—the first at least as a consequence of the second. I shall here confine myself to showing that the third can be easily proved, without invoking anything that was not well known to Fermat himself.

The system of equations (1) $x^2 + 2y^2 = z_1^2$ *and* (2) $x^2 + y^2 = z_2^2$ *is impossible in integers,* $y > 0$. On subtracting the second from the first, and transposing, we get (3) $z_2^2 + y^2 = z_1^2$ as a necessary complement to the system.† In the first of these equations it is evident from the formulae already given, that in primitive solutions y must be the even number;‡ and therefore y must be the even number in the others also. Now the formulae, known to Fermat, that will yield all primitive solutions (for it is sufficient to prove the proposition of these), are

of (1)	$x = 2p_1^2 - q_1^2$,	$y = 2p_1 q_1$,	$z_1 = 2p_1^2 + q_1^2$,§
of (2)	$x = p_2^2 - q_2^2$,	$y = 2p_2 q_2$,	$z_2 = p_2^2 + q_2^2$,
of (3)	$z_2 = p_3^2 - q_3^2$,	$y = 2p_3 q_3$,	$z_1 = p_3^2 + q_3^2$,

where the p's and q's, in their respective formulae, are prime to each other, and there can be no interchange of the formulae of x and y. As y is the same in all, we have $p_1 q_1 = p_2 q_2 = p_3 q_3$. Comparing (1) and (2), we see that if $p_1 = p_2$ and $q_1 = q_2$, x would be greater in (1) than in (2); and

---

* *Op. cit.*, p. 117.

† Also, because of the common values of y and of x in the first two, we get

$$(4)\quad z_1^2 - 2z_2^2 = 3x^2 \qquad \text{and} \qquad (5)\quad z_1^2 - z_2^2 = 3y^2.$$

Euler, in his *Algebra*, II, ii, § 71, has in a very simple manner proved that (4) is impossible in integers. We might stop at this as proving the impossibility of the simultaneous existence of (1) and (2) in integers. But as it was unknown to Fermat, I proceed without making use of it.

‡ Even if it were not so in general, it would have to be so here, because in primitive solutions of (1) and (2) z_1 and z_2 must be odd, and therefore in (3) y must be even. Or, on adding (1) and (2), we get $z_1^2 + z_2^2 = 2x^2 + 3y^2$, in which the first member is even, and so also the second; which requires y to be even.

§ This is the alternative set of formula noticed above. It is used here merely as the more convenient. The other would work out the same, with only some interchanges of letters and some alterations of signs, all immaterial.

it would be still more so if $p_1 > p_2$ and $q_1 < q_2$; wherefore the only remaining possibility is $p_1\sqrt{2} > p_2$ but $p_1 < p_2$ and $q_1 > q_2$. And comparing (2) and (3), we see that to have z_2 the same in both, we must have $p_3 > p_2$ and $q_3 < q_2$. Thus we have

$$p_3 > p_2 > p_1 \quad \text{and} \quad q_1 > q_2 > q_3.$$

Therefore the proportions

$$p_3 : p_2 = q_2 : q_3 \quad \text{and} \quad p_3 : p_1 = q_1 : q_3$$

are in descending order, or $p_3 : p_2 = \frac{1}{m} p_3 : \frac{1}{m} p_2$, which shows that $q_2 = \frac{1}{m} p_3$, and $q_3 = \frac{1}{m} p_2$ or $p_2 = m q_3$; and $p_3 : p_1 = \frac{1}{n} p_3 = \frac{1}{n} p_1$, which shows that $q_1 = \frac{1}{n} p_3$, and $q_3 = \frac{1}{n} p_1$ or $p_1 = n q_3$, m and n either integers or fractions > 1, and $m > n$ because q_2 is $< q_1$ and p_2 is $> p_1$. We can now put all the symbols for the generators in terms of p_3 and q_3, omitting the subscripts as no longer needed. The solutions become

of (1) $\quad x = 2(nq)^2 - \left(\frac{p}{n}\right)^2, \quad y = 2nq\frac{1}{n}p, \quad z_1 = 2(nq)^2 + \left(\frac{p}{n}\right)^2,$

of (2) $\quad x = (mq)^2 - \left(\frac{p}{m}\right)^2, \quad y = 2mq\frac{1}{m}p, \quad z_2 = (mq)^2 + \left(\frac{p}{m}\right)^2,$

of (3) $\quad z_2 = p^2 - q^2, \quad y = 2pq, \quad z_1 = p^2 + q^2.$

Equating the two formulae of x, the two of z_1, and the two of z_2, and seeking the value of p in q, we get respectively

$$p = mnq\sqrt{\frac{2n^2 - m^2}{m^2 - n^2}}, \quad p = nq\sqrt{\frac{2n^2 - 1}{n^2 - 1}}, \quad p = mq\sqrt{\frac{m^2 + 1}{m^2 - 1}},$$

which, inserted in both the formulae of x, of z_1, and of z_2 respectively, yield in each case the same result, as follows:

$$x = q^2\left(\frac{m^4 - 2n^4}{m^2 - n^2}\right), \quad z_1 = q^2\left(\frac{2n^4 - 1}{n^2 - 1}\right), \quad z_2 = q^2\left(\frac{m^4 + 1}{n^2 - 1}\right).$$

Now, instead of questioning whether it is possible to find single values for m and n that will make x, z_1 and z_2 integral, we may observe that for y to be the same in the three equations p must be the same in its three formulae. Let us, then, inquire whether, or how, this is possible. We may equate the formulae of p in pairs. In each equation we may eliminate q^2, square the formulae, clear the fractions, and then put either

m or $n = 1$, to see what ratio the other has to it. In each case we find the other $= 1$, that is, $m = n$. But above we have seen that m must be $> n$. Thus the solution of this system in integers involves contradictory requirements. Therefore the system cannot be solved in integers.

Consequently the hypothesis of $\frac{1}{2}xy$ in $x^2+y^2 = z^2$ in integers being a square, which involves this impossible system, is also impossible.

Another obvious conclusion is that, since x^2, x^2+y^2, x^2+2y^2, which are in arithmetical progression with difference y^2, cannot all three be squares, no arithmetical progression of three integers all squares and difference a square is possible. This legitimizes Leibniz's proof of the impossibility of $\frac{1}{2}xy$ in $x^2+y^2 = z^2$ being a square. Leibniz proved, in a Corollary, the impossibility of $x^4 - y^4$ being a square. This legitimizes Wallis's proof.

The truth is, many impossibilities about single equations can be proved as deductions from the impossibility of systems of equations. This method of proof, however, was never employed by Fermat. I say this in spite of his use of the system $u^2 \pm v^2 = \square$ in the very Note before us, because there he showed the impossibility of this system only as involved in the rejected supposition: he did not prove its impossibility in general; nor was its impossibility essential to his argument, which used only the descent of $u^2+v^2 = \square$. And nowhere else in his extant writings did he make use of any impossible system. Nor, so far as I know, did Euler employ this method, nor Legendre. It seems to have been employed first by Barlow.

A NEW DERIVATION OF THE LORENTZ TRANSFORMATION.*

By P. Y. Chou.

The characteristics of an electromagnetic wave are described by a wave function, U, satisfying the equation,

$$\Box U \equiv \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = 0, \tag{1}$$

in which c is the velocity of propagation and (x, y, z, t) are the space and time coördinates used by an observer in his reference system, S. Let (x', y', z', t') be the coördinates used by a second observer to record the same phenomenon observed in another system, S', which travels with a uniform linear velocity, v, in S. Since all experimental tests have failed to determine which one of the two systems S and S' is more advantageous for a simple description of electromagnetic phenomena or to lay any restriction on v, we lay down the fundamental postulate that there is a set including an infinite number of such systems all of which are equally advantageous and that the kinematical relations between the coördinates (x, y, z, t) and (x', y', z', t') in two members of this set leave the wave equation, (1), invariant. Symbolically this means that

$$\Box U \equiv \lambda(x, y, z, t, c, v)\Box' U, \qquad (\lambda \neq 0) \tag{2}$$

where U is assumed to be a scalar in this transformation, λ, an unknown function of the arguments indicated and

$$\Box' \equiv \frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2} - \frac{1}{c'^2}\frac{\partial^2}{\partial t'^2}. \tag{3}$$

The quantity, c', is the velocity of propagation of plane electromagnetic waves (supposed to be observed) in S' and might *a priori* be a function of c and v. It will be seen presently that the following relations necessarily hold

* Received July 19, 1927, in revised form January 3, 1928.

$$c' = c, \qquad \lambda(x, y, z, t, c, v) \equiv 1,$$

$$(4) \qquad x' = \varkappa(x - vt), \quad y' = y, \quad z' = z, \quad t' = \varkappa\left(t - \frac{vx}{c^2}\right),$$

$$\varkappa = [1 - v^2/c^2]^{-1/2}$$

provided the identity, (2), and the following postulates and boundary conditions are satisfied:

A. The functions of transformation,

$$(5) \qquad x' = x'(x, y, z, t), \quad y' = y'(x, y, z, t), \quad z' = z'(x, y, z, t), \quad t' = t'(x, y, z, t),$$

are real, single valued, continuous and have continuous second partial derivatives in the corresponding real domains of S and S'.

B. They are all independent of one another and all increase in the same sense as the corresponding quantities (x, y, z, t) so that the Jacobian satisfies the condition,

$$(6) \qquad J \equiv \frac{\partial(x', y', z', t')}{\partial(x, y, z, t)} > 0$$

everywhere.

C. Axes OY, OZ of S are parallel to $O'Y'$ and $O'Z'$ of S' respectively and OX coincides with $O'X'$. S' travels with a uniform velocity, v, in the positive direction of OX in S.

D. In S and S' the same units of length and time measurements are adopted.

E. The boundary conditions are

$$(7) \qquad \begin{aligned} x' = t' &= 0, && \text{when } x = t = 0; \\ y' &= 0, && \text{when } y = 0; \\ z' &= 0, && \text{when } z = 0. \end{aligned}$$

To begin our proof we shall show that $\lambda(x, y, z, t, c, v)$ is necessarily unity. Let S'' be another reference system that travels with a uniform velocity, $v + u$, (observed by S) along OX in S. Then it will have a velocity, v', observed by S' and v' will be a function of v, u and possibly c. By (2) we have the relations

$$(8) \qquad \begin{gathered} \Box' U = \lambda(x', y', z', t', c', v')\, \Box'' U, \\ \Box U = \lambda(x, y, z, t, c, v)\, \Box' U = \lambda(x, \cdots, t, c, v)\, \lambda(x', \cdots, t', c', v')\, \Box'' U, \end{gathered}$$

and if we transform our coördinates from S to S'' directly, we get

$$(9) \qquad \Box U = \lambda(x, y, z, t, c, v + u)\, \Box'' U.$$

Comparing the coefficients of $\Box''U$ in (8) and (9) we get the identity,

$$\lambda(x, y, z, t, c, v)\,\lambda(x', y', z', t', c', v') \equiv \lambda(x, y, z, t, c, v+u), \tag{10}$$

from which we can derive several important consequences.

In the first place let $u = 0$. Then S' and S'' will travel with the same velocity, v, in S and $v' = 0$. Hence (10) becomes

$$\lambda(x', y', z', t', c', 0) = 1 \tag{11}$$

where (x', y', z', t') may have any arbitrary values.

Next let $v = 0$. Then S' will be stationary in S. Since S and S' use the same units of measurements, $c' = c$ and S'' will be observed to have the same velocity, u, by both S and S'. By (11), (10) becomes

$$\lambda(x', y', z', t', c', u) = \lambda(x, y, z, t, c, u). \tag{12}$$

But (x', y', z', t') can differ from the corresponding (x, y, z, t) by any finite constants and since (x, y, z, t) are all independent variables, relation (12) is absurd unless λ is independent of (x, y, z, t). Consequently

$$\lambda(x, y, z, t, c, v) \equiv \lambda(c, v)$$

and (10) and (11) can be rewritten in the form,

$$\lambda(c, v)\,\lambda(c', v') \equiv \lambda(c, v+u), \qquad \lambda(c', 0) = 1. \tag{13}$$

The left hand side of the first equation of (13) remains unaltered if we interchange c, v and c', v' respectively. The right hand side becomes $\lambda(c', v'+u)$. Hence we get a new identity

$$\lambda(c, v+u) \equiv \lambda(c', v'+u). \tag{14}$$

Putting $u = 0$ in (14) and since $v' = 0$, when $u = 0$, we obtain from (13) and (14) the relation,

$$\lambda(c, v) = 1. \tag{15}$$

In the foregoing consideration no restriction has been laid upon c and v, so (15) holds for all c and v and is therefore an identity.

Consider a point (x_0, y, z) in S at the instant $t = 0$. The corresponding x' coördinate of the point in S' according to (5) will be

$$x' = x'(x_0, y, z, 0). \tag{16}$$

The function, x', in (16) will remain unaltered if we make a simultaneous linear translation of the point (x_0, y, z) and the coördinate system, S', along OX to a distance vt in S. This amounts to shifting the origin O of S to a distance, $-vt$, along the X-axis. In short we have the identity,*

$$x'(x_0, y, z, 0) \equiv x'(x_0 + vt, y, z, t). \tag{17}$$

Put $x_0 = x - vt$ in (17) and get $x'(x, y, z, t) = x'(x - vt, y, z, 0)$. Likewise we obtain similar expressions of y' und z'. The most general equations of transformation (5) then appear to be

$$\begin{gathered} x' = x'(x - vt, y, z), \quad y' = y'(x - vt, y, z), \quad z' = z'(x - vt, y, z), \\ t' = t'(x, y, z, t). \end{gathered} \tag{18}$$

Let $x = x_1$, $y = x_2$, $z = x_3$, $t = x_4$. Unless stated otherwise we adopt the summation convention of Tensor Analysis in such a way that when two same Greek letters are present in one term, the sum extends over 1, 2, 3, 4 and when two Latin letters occur, the sum extends over 1, 2, 3. For example

$$\frac{\partial U}{\partial x_i} = \frac{\partial U}{\partial x'_\alpha}\frac{\partial x'_\alpha}{\partial x_i} = \frac{\partial U}{\partial x'_1}\frac{\partial x'_1}{\partial x_i} + \frac{\partial U}{\partial x'_2}\frac{\partial x'_2}{\partial x_i} + \frac{\partial U}{\partial x'_3}\frac{\partial x'_3}{\partial x_i} + \frac{\partial U}{\partial x'_4}\frac{\partial x'_4}{\partial x_i},$$

$$\alpha_{ik}\alpha_{jk} = \alpha_{i1}\alpha_{j1} + \alpha_{i2}\alpha_{j2} + \alpha_{i3}\alpha_{j3}.$$

With the above agreement and by introducing the fundamental postulate (2) we obtain

$$\frac{\partial U}{\partial x_i} = \frac{\partial U}{\partial x'_\alpha}\frac{\partial x'_\alpha}{\partial x_i}; \quad \frac{\partial^2 U}{\partial x_j\,\partial x'_i} = \frac{\partial^2 U}{\partial x'_\beta\,\partial x'_\alpha}\frac{\partial x'_\beta}{\partial x_j}\frac{\partial x'_\alpha}{\partial x_i} + \frac{\partial U}{\partial x'_\alpha}\frac{\partial^2 x'_\alpha}{\partial x_j\,\partial x_i},$$

$$\begin{aligned} \Box U = {} & \frac{\partial^2 U}{\partial x'_\beta\,\partial x'_\alpha}\frac{\partial x'_\beta}{\partial x_j}\frac{\partial x'_\alpha}{\partial x_j} + \frac{\partial U}{\partial x'_\alpha}\frac{\partial^2 x'_\alpha}{\partial x_j\,\partial x_j} \\ & - \frac{1}{c^2}\left[\frac{\partial^2 U}{\partial x'_\beta\,\partial x'_\alpha}\frac{\partial x'_\beta}{\partial x_4}\frac{\partial x'_\alpha}{\partial x_4} + \frac{\partial U}{\partial x'_\alpha}\frac{\partial^2 x'_\alpha}{\partial x_4^2}\right] \\ \equiv {} & \Box' U \equiv \frac{\partial^2 U}{\partial x'_i\,\partial x'_i} - \frac{1}{c'^2}\frac{\partial^2 U}{\partial x_4'^2}. \end{aligned} \tag{19}$$

Equating the coefficients of the derivatives of U on both sides of (19) we get the following sets of differential equations,

* The original idea of this consideration was due to Prof. G. A. Bliss when he gave the derivation of Lorentz transformation in his course, "Calculus of Variations and its Application to Problems in Mathematical Physics" in Autumn 1926 at the University of Chicago.

$$(20)\qquad \frac{\partial x_i'}{\partial x_j}\frac{\partial x_k'}{\partial x_j}-\frac{1}{c^2}\frac{\partial x_i'}{\partial x_4}\frac{\partial x_k'}{\partial x_4}=\begin{cases}1 & (i=k),\\ 0 & (i\neq k),\end{cases}\qquad (i,k=1,2,3);$$

$$(21)\qquad \frac{\partial x_i'}{\partial x_j}\frac{\partial x_4'}{\partial x_j}-\frac{1}{c^2}\frac{\partial x_i'}{\partial x_4}\frac{\partial x_4'}{\partial x_4}=0,\qquad (i=1,2,3);$$

$$(22)\qquad \frac{\partial x_4'}{\partial x_j}\frac{\partial x_4'}{\partial x_j}-\frac{1}{c^2}\left(\frac{\partial x_4'}{\partial x_4}\right)^2=-\frac{1}{c'^2};$$

$$(23)\qquad \frac{\partial^2 x_\alpha'}{\partial x_j\,\partial x_j}-\frac{1}{c^2}\frac{\partial^2 x_\alpha'}{\partial x_4^2}=0,\qquad (\alpha=1,2,3,4).$$

Now consider the first set of (20) where $i=k$. By (18) we have

$$(24)\qquad \frac{\partial x_i'}{\partial x_1}=\varkappa\frac{\partial x_i'}{\partial w},\qquad \frac{\partial x_i'}{\partial x_4}=-\varkappa v\frac{\partial x_i'}{\partial w},$$

where $w=\varkappa(x-vt)$, $\varkappa=[1-v^2/c^2]^{-1/2}$. Hence (20) becomes

$$(25)\qquad \left(\frac{\partial x_i'}{\partial w}\right)^2+\left(\frac{\partial x_i'}{\partial y}\right)^2+\left(\frac{\partial x_i'}{\partial z}\right)^2=1,\qquad (i=1,2,3).$$

One set of complete solutions of (25) is

$$(26)\qquad x_i'=\alpha_{i1}w+\alpha_{i2}x_2+\alpha_{i3}x_3+\alpha_i,\qquad (i=1,2,3),$$

where $\alpha_{i1}^2+\alpha_{i2}^2+\alpha_{i3}^2=1$, $\alpha_i=$ additive constants.

The second set of (20) when $i\neq k$ gives the further restriction on α_{ik},

$$(27)\qquad \alpha_{ij}\,\alpha_{kj}=0.$$

The above analysis shows that the system $S'(x_1', x_2', x_3')$ is orthogonal and the α_{ik}'s are the direction cosines of the axes $O'X'$, $O'Y'$ and $O'Z'$ in the system $S(x_1, x_2, x_3)$. But (26) must satisfy the conditions (6) and (7) laid down at the outset. Hence

$$(28)\qquad \alpha_{ik}=\begin{cases}1 & (i=k),\\ 0 & (i\neq k),\end{cases}\qquad \alpha_i=0\qquad (i=1,2,3),$$

$$(29)\qquad x'=w=\varkappa(x-vt),\qquad y'=y,\qquad z'=z.$$

Solutions of x', y', z' given in (29) are unique, because by going through a similar process as (24) the first three equations of (23) become Laplace's equations in the (w, y, z) space, viz.,

$$(30)\qquad \frac{\partial^2 x_i'}{\partial w^2}+\frac{\partial^2 x_i'}{\partial y^2}+\frac{\partial^2 x_i'}{\partial z^2}=0,\qquad (i=1,2,3)$$

and any other solutions of (30) that must also satisfy (20) and (21) and vice versa with the boundary conditions (7) will be necessarily identical with (29).

Next we come to prove the important consequence of the present discussion that the velocities of light c and c' in S and S' respectively are numerically equal. Starting with this consequence as a fundamental postulate Einstein derived the Lorentz transformation in 1905. So far we obtained our results (29) by considering the transformation from $\square U$ to $\square' U$ in (19). On the other hand we can transform our wave equation from S' to S and go through the precise reasoning from (19) to (29). All the postulates based upon which we obtain (29) are reciprocal in nature and can be equally well applied to the unknown functions (x, y, z, t). The set of unique solutions that correspond to (29) will be then

$$x = \varkappa'(x' + v t'), \quad y = y', \quad z = z', \tag{31}$$

where $+vt'$ is used because by hypothesis, D, S' and S have the same units of measurements, so S is observed to travel with a uniform velocity, $-v$, along $O'X'$ in S'. Furthermore we write

$$\varkappa' = [1 - v^2/c'^2]^{-1/2}. \tag{32}$$

Eliminating x' from (29) and (31) we get

$$t' = \varkappa t - \frac{1}{\varkappa' v}(\varkappa'\varkappa - 1)x. \tag{33}$$

The above expression for t' must satisfy (22). After some algebraic operations we finally land at the relation between the two velocities c and c',

$$c' = \pm c, \text{ and } \varkappa' = \varkappa = [1 - v^2/c^2]^{-1/2}. \tag{34}$$

The negative sign in (34) is excluded by (6) because by (6) if a ray of light travels along the positive OX-axis in S, it goes also along the positive branch of $O'X'$ in S'. The last equation of Lorentz transformation becomes

$$t' = \varkappa\left(t - \frac{vx}{c^2}\right) \tag{35}$$

which is also a solution of (21) and the fourth equation of (23). Consequently every one of the fourteen differential equations from (20) to (23) is satisfied by the present solution given in (29) and (35).

The foregoing derivation of the Lorentz transformation from the consideration of the invariance of the wave equation seems to fit the fundamental

postulate of the general theory of relativity that physical laws under the disguise of differential equations should remain covariant under all point transformations more properly than the original derivation from the hypothesis that the velocity of light is a universal constant in both S and S'. Here we notice that the invariance of the metric,

$$ds^2 = c^2 dt^2 - (dx^2 + dy^2 + dz^2), \tag{36}$$

follows as a consequence of the invariance of $\Box U$. The assumption that U is a scalar under the present coördinate transformation is still an enigma. Its justification can be only obtained from experimental physicists who are primarily concerned with observable quantities like frequency, intensity, velocity and phase difference of an electromagnetic wave all of which are concisely stated by the wave function, U.

In fine I wish to express my sincere gratitude to Prof H. Bateman whose deep-sighted citicism has decided many subtle points within this paper and Mr. B. Podolsky for discussions on some physical conceptions which underlie this investigation.

California Institute of Technology,
Pasadena, Calif.

ON m-DIMENSIONAL CROSS-RATIOS.*

By Louis Weisner.†

I. Introduction.

In view of the fundamental place which the cross-ratio of four collinear points occupies in Projective Geometry, it is not surprising that attempts have been made to generalize this concept.‡

All of the proposed generalizations have one property in common: they are invariant under projection. But most of them do not possess other properties which are generalizations of the known properties of the cross-ratios of four points. The reason is that the characteristic property of the cross-ratio is not its invariance under projection—as an arbitrary function of the cross-ratio also possesses this property—but is that expressed by the theorem: A necessary and sufficient condition that two sets of four collinear points be projective is that their respective cross-ratios be equal. In generalizing to m dimensions we seek functions of $m+3$ independent points in S_m which furnish necessary and sufficient conditions that two sets of $m+3$ independent points of S_m be projective. Such functions (m of them in S_m) have been suggested by Veblen.§ The present paper is devoted to the consideration of these cross-ratios. It will appear that they have properties which are generalizations of the properties of the cross-ratio of four points.

Some of the subsequent results have been obtained synthetically by G. Kohn,‖ who generalized von Staudt's theory of throws to m dimensions.

* Received June 22, 1927. Presented to the American Mathematical Society, May 7, 1927.

† National research fellow.

‡ A. F. Möbius, *Werke*, vol. 1, p. 462; E. Heis, *Stereometrische Sätze etc.*, Grunert Archiv, vol. 31, p. 37; W. K. Clifford, *General theory of anharmonics*, Proc. Lond. Math. Soc., vol. 2 (1866), p. 3, Math. Papers, p. 110; C. W. Merrifield, *Anharmonic ratio of a pencil of five lines in space*, Proc. Lond. Math. Soc., vol. 5 (1876), p. 94; W. J. C. Sharp, *Simplicissima of n dimensions*, Proc. Lond. Math. Soc., vol. 19 (1888), p. 433; J. Brill, *On certain analogues of anharmonic ratios*, Quart. Jour. Math., vol. 29 (1898), p. 286; E. O. Lovett, *General theory of anharmonics*, Proc. Lond. Math. Soc., vol. 29 (1898), p. 566; Veblen and Young, *Projective Geometry*, vol. 2, p. 55; A. B. Coble, *Point sets and Cremona groups, Part 1*, Trans. Amer. Math. Soc., vol. 16, p. 187.

§ L. c.

‖ G. Kohn, *Über die Erweiterung eines Grundbegriffes der Geometrie der Lage*, Math. Annalen, vol. 46, p. 285.

II. Properties of cross-ratios.

1. We are concerned with $m+3$ ordered points $P_i(x_{i1}, x_{i2}, \cdots, x_{i,m+1})$, $(i = 1, \cdots, m+3)$ in S_m. With the notation

$$(12\cdots m\,m+1) \equiv (P_1 P_2 \cdots P_m P_{m+1}) \equiv \begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1,m+1} \\ x_{21} & x_{22} & \cdots & x_{2,m+1} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m+1,1} & x_{m+1,2} & \cdots & x_{m+1,m+1} \end{vmatrix}$$

we define Veblen's cross-ratios as follows:

$$r_{1m}\,(P_1 P_2 \cdots P_{m+2} P_{m+3}) \equiv \frac{(123\cdots m\,m+2)\,(m+323\cdots m\,m+1)}{(123\cdots m\,m+3)\,(m+223\cdots m\,m+1)},$$

$$r_{2m}\,(P_1 P_2 \cdots P_{m+2} P_{m+3}) \equiv \frac{(123\cdots m\,m+2)\,(1\,m+33\cdots m\,m+1)}{(123\cdots m\,m+3)\,(1\,m+23\cdots m\,m+1)},$$

$$\cdots\cdots\cdots

2. Necessary and sufficient conditions that the $m+3$ independent points $P_1, \cdots, P_{m+3}$ be projective with the $m+3$ independents points $P_1', \cdots, P_{m+3}'$ respectively, are that*

$$r_{im}(P_1\,P_2\cdots P_{m+2}\,P_{m+3}) = r_{im}(P_1'\,P_2'\cdots P_{m+2}'\,P_{m+3}'), \qquad (i=1,\cdots,m).$$

3. The collineations which have the $m+1$ independent fixed points $P_1, \cdots, P_{m+1}$ may be obtained by solving the equations

$$r_{im}(P_1\,P_2\cdots P_{m+1}\,xx') = k_i \neq 0, \qquad (i=1,\cdots m)$$

for $x_1' : x_2' : \cdots : x_{m+1}'$.

4. If a collineation of S_m has $m+1$ independent fixed points, the cross-ratios of the fixed points and any two corresponding points are constants.†

5. Necessary and sufficient conditions that $m+3$ points of S_m be dependent are that two of their standard cross-ratios be equal or that one of them have the value 0, 1, ∞, or 0/0.

6. The locus of a point which moves in S_m so that the cross-ratios of the lines joining it to $m+2$ fixed independent points is constant is a rational norm-curve passing through the fixed points.

III. Associated point sets.

4. Given $n \geqq m+3$ independent points $P_1, \cdots, P_{m+2}, P_1', \cdots, P_k'$ in S_m, $(n = m+k+2)$, we form the k sets of cross-ratios

$$\begin{aligned}
&r_{1m}(P_1\,P_2\cdots P_{m+2}\,P_1'), \quad \cdots, \quad r_{mm}(P_1\,P_2\cdots P_{m+2}\,P_1'),\\
&r_{1m}(P_1\,P_2\cdots P_{m+2}\,P_2'), \quad \cdots, \quad r_{mm}(P_1\,P_2\cdots P_{m+2}\,P_2'),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&r_{1m}(P_1\,P_2\cdots P_{m+2}\,P_k'), \quad \cdots, \quad r_{mm}(P_1\,P_2\cdots P_{m+2}\,P_k').
\end{aligned}$$

Similarly, with n independent points $Q_1, \cdots, Q_{k+2}, Q_1', \cdots, Q_m'$ in S_k we form the m sets of cross-ratios

$$\begin{aligned}
&r_{1k}(Q_1\,Q_2\cdots Q_{k+2}\,Q_1'), \quad \cdots, \quad r_{kk}(Q_1\,Q_2\cdots Q_{k+2}\,Q_1'),\\
&r_{1k}(Q_1\,Q_2\cdots Q_{k+2}\,Q_2'), \quad \cdots, \quad r_{kk}(Q_1\,Q_2\cdots Q_{k+2}\,Q_2'),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&r_{1k}(Q_1\,Q_2\cdots Q_{k+2}\,Q_m'), \quad \cdots, \quad r_{kk}(Q_1\,Q_2\cdots Q_{k+2}\,Q_m').
\end{aligned}$$

---

* As an application we may give a simple analytic proof of the following theorem: The plane 5-point obtained by joining the alternate vertices of a given 5-point is projective with the given 5-point. For a synthetic proof, see Clebsch, *Ueber das ebene Fünfeck*, Math. Annalen, vol. 4, p. 477.

† Kohn (l. c., p. 291, 325) calls these constants the *throw* of the collineation.

We choose $Q_1', \cdots, Q_m'$ so that

$$r_{jk}(Q_1\, Q_2 \cdots Q_{k+2}\, Q_i') = r_{im}(P_1\, P_2 \cdots P_{m+2}\, P_j'), \quad \begin{pmatrix} i = 1, \cdots, m \\ j = 1, \cdots, k \end{pmatrix},$$

or, more fully,

$$\frac{(Q_1\, Q_2 \cdots Q_k\, Q_{k+2})\,(Q_1\, Q_2 \cdots Q_{j-1}\, Q_i' \quad Q_{j+1} \cdots Q_k\, Q_{k+1})}{(Q_1\, Q_2 \cdots Q_k\, Q_i' \quad)\,(Q_1\, Q_2 \cdots Q_{j-1}\, Q_{k+2}\, Q_{j+1} \cdots Q_k\, Q_{k+1})}$$
$$= \frac{(P_1\, P_2 \cdots P_m\, P_{m+2})\,(P_1\, P_2 \cdots P_{i-1}\, P_j' \quad P_{i+1} \cdots P_m\, P_{m+1})}{(P_1\, P_2 \cdots P_m\, P_j' \quad)\,(P_1\, P_2 \cdots P_{i-1}\, P_{m+2} P_{i+1} \cdots P_m\, P_{m+1})}.$$

The rows of one of the above rectangular arrays are then equal to the columns of the other.

Two sets of n points, one (P_n^m) in S_m, the other (Q_n^k) in S_k, which are so related, shall be called *associated* point sets. They are associated in a definite order: P_i with Q_i' $(i = 1, \cdots, m)$; P_i' with Q_i $(i = 1, \cdots, k)$; P_{m+1} with Q_{k+2}; P_{m+2} with Q_{k+1}. The preceding definition is readily reconciled with Coble's* and has the advantage of enabling one to determine by direct calculations whether or not two sets of points are associated.

Construct the nth order determinant D, each row of which consists of the ordered coördinates of one of the points P_n^m followed by the ordered coördinates of the associated point of Q_n^k. The left member of the last equation may be obtained from the right by replacing each determinant, regarded as a minor of D, by its algebraic complement; and the right member may be obtained from the left similarly. This rule is important in deriving projective invariants of Q_n^k from those of P_n^m, the equation referred to being an example.

5. If a permutation of the points P_n^m changes $r_{im}(P_1\, P_2 \cdots P_{m+2}\, P_j')$ to $r_{im}'(P_1\, P_2 \cdots P_{m+2}\, P_j')$, the latter equals a rational function of $r_{im}(P_1\, P_2 \cdots P_{m+2}\, P_j')$, $(i = 1, \cdots, m;\ j = 1, \cdots, k)$, whose numerator and denominator are at most of degree 2 in the arguments. The $n!$ permutations of the points give rise to $n!$ transformations which form a Cremona group $G_{n!}$ of order $n!$ in km variables, which, for $k = 1$, reduces to E. H. Moore's cross-ratio group.† This group has been studied by Coble.‡ We note here a few properties which are immediate consequences of the properties of cross-ratios.

Because of the relations between the cross-ratios of two associated point sets, it is evident that associated point sets give rise to the same Cremona group. Regarding the mk cross-ratios as non-homogeneous coördinates of

* L. c., p. 156.

† Amer. Jour. Math., vol. 22, p. 288.

‡ L. c., p. 182.

a point R in S_{mk}, we have a representation of the class of all ordered point sets in S_m projective with P_n^m and of the class of all ordered point sets in S_k projective with Q_n^k, by the point R in S_{mk}.* In general, this point has $n!$ conjugates under $G_{n!}$, each of which is the representation in S_{mk} of the sets of points obtained from P_n^m or Q_n^k by some permutation of the points. But a particular point R_μ of S_{mk} may have $\mu < n!$ conjugates. This occurs when, and only when, $n!/\mu$ transformations of $G_{n!}$ leave R_μ fixed. We infer from Theorem 2 of § 3 that there are collineations S_m which effect those permutations of the points P_n^m that induce the transformations of $G_{n!}$ which leave R_μ fixed. Knowing the points of S_{mk} which have less than $n!$ conjugates under $G_{n!}$, one could determine all sets of n points in S_m and in S_k which are self-projective in some order other than the identical order. The case $n = m + 3$, $k = 1$, has been discussed by E. H. Moore.†

6. An absolute projective invariant I of P_n^m assumes the same value for two projective P_n^m's. Hence, if the mk cross-ratios $r_{im}(P_1 P_2 \cdots P_{m+2} P_j')$ of § 4 are given, the value of I is determined. It follows that I equals a function I' of these cross-ratios. If I is symmetric in the points P_n^m, I' is invariant under the Cremona $G_{n!}$. Conversely, from every absolute invariant I' of $G_{n!}$ we obtain an absolute projective symmetric invariant I of P_n^m by simply replacing each argument of I' by its value in terms of the coördinates of the points P_n^m.

Starting with I, let us form I', and from I' form the invariant I'' of the associated point set Q_n^k. As associated point sets give rise to the same $G_{n!}$, the remarks of the preceding paragraph assure us that I'' is symmetric in the points Q_n^k. Similarly, from I'' we obtain I. We may pass directly from I to I'', or from I'' to I, without the intervention of I', by replacing each determinant by its complement (§ 4). There is thus established a one-one correspondence between the *absolute* invariants of associated point sets.

But such a correspondence cannot be established in the same way between the *relative* invariants of associated point sets, as Coble contends,‡ because the function so derived from a symmetric function of P_n^m may be an *alternating* function of Q_n^k. I have verified that this actually occurs in the case of the I_{18} of the binary quintic, using for this purpose the expression for I_{18} in terms of the roots of the quintic given in Faa di Bruno's *Theorie der Binaren*, p. 327.

* L. c., p. 176.

† L. c., p. 291.

‡ L. c., p. 166.

HARVARD UNIVERSITY.

TWISTED CURVES CLASSIFIED AS TO THEIR OSCULATING AND QUASI-OSCULATING SPHERES.*

By Arthur Ranum.

1. The quasi-osculating sphere of a twisted curve at a point, discovered by Hostinsky and myself in 1909,† is dual to the well-known osculating sphere. Hence it may be briefly defined as a sphere touching four consecutive osculators‡ (osculating planes) of the curve. In a paper soon to appear in the Mathematische Annalen§ I have studied three new classes of curves, as follows:

Class (K). Every curve C of this class is the curve of centers of a family of spheres osculating a curve C_1 and at the same time quasi-osculating a curve C_2; C is a twisted parabola.

Class (K_1), consisting of the curves C_1 just mentioned, each of which is osculated by the quasi-osculating spheres of another curve C_2.

Class (K_2), consisting of the curves C_2, each of which is quasi-osculated by the osculating spheres of a curve C_1.

Let us now form a table including some of the simpler known classes of curves together with these three new classes. Each class will be defined by an intrinsic equation involving one or more of the six variables‖

$$\begin{array}{ccc} s & \theta & \eta \\ & \rho \quad \sigma & \\ & \tau & \end{array}$$

where s is the arc-length, $1/\rho$ and $1/\tau$ are the curvature and torsion, respectively, $\theta = \int ds/\rho$, $\eta = \int ds/\tau$, and $\sigma = \rho/\tau$; s is dual to η and ρ to σ, while θ and τ are self-dual.

The simplest of all classes are evidently those for which τ, ρ, or σ is constant; we call them (I), (I_1), and (I_2), respectively. The next in order of simplicity, from the stand-point of sphere-geometry, are the class (J_1)

* Received January 6, 1928.

† Hostinsky, Journ. de Math., (6), vol. 5, p. 285. Ranum, Bull. Amer. Math. Soc., vol. 17, pp. 67–68; Quart. Journ. Math., vol. 46 (1915), pp. 364–366.

‡ G. Loria, Curve Sghembe Speciali (1921), vol. 1, p. 2.

§ Entitled "Spheres osculating a curve and quasi-osculating another curve". This paper will be referred to hereafter as M. A.

‖ See M. A., § 1.

of curves having a fixed osculating sphere (spherical curves) and the dual class (J_2) of curves C having a fixed quasi-osculating sphere S (which we shall call quasi-spherical curves); in the latter case the tangents and osculators to C all touch S, and C is a geodesic on a cone whose vertex is the center of S.

The intrinsic equations of the curves of the classes (J_1) and (J_2) are given by G. Loria, loc. cit., vol. 2, p. 55, equation (35) and p. 211, equation (17); and for the classes (K_1), (K_2) and (K) the intrinsic equations are given in M. A., § 6, equation (16), § 11, equation (26), and § 4, equation (5).

2. The results may be tabulated as follows:

(I) $\tau = c$. Curve of constant torsion.

(I_1) $\rho = c$. Twisted circle (of constant curvature).

(J_1) $\rho = c\cos\eta$. Spherical curve (having a fixed osculating sphere of radius c).

(K_1) $\rho = c\sec\eta$. Curve osculated by the quasi-osculating spheres of another curve.

(K) $\rho = 2c\sec^3\theta$ $(\tau \neq 0)$. Twisted parabola.

(I_2) $\sigma = c$. Cylindrical helix (of constant slope).

(J_2) $\sigma = s/c$. Quasi-spherical curve (having a fixed quasi-osculating sphere of radius c).

(K_2) $\sigma^{-1} - \sigma = s/c$. Curve quasi-osculated by the osculating spheres of another curve.

3. (I_1) is dual to (I_2), (J_1) to (J_2), and (K_1) to (K_2), while (I) is self-dual. Obviously (K) has no dual class. This tabulation emphasizes two distinct principles of unity and systematic arrangement: first, the principle of duality, and second, the connection between the three classes I_1, J_1, K_1 (and between I_2, J_2, K_2), due to the fact that their equations involve the same variables ρ, η (σ, s) in a somewhat similar manner.

Indeed, consider two curves C, C' belonging to the respective classes J_1, K_1 and therefore satisfying equations of the form $\rho = c\cos\eta$, $\rho' = c'\sec\eta'$. Let a point at which ρ has its maximum value c be called a vertex of C; and similarly let a point at which ρ' has its minimum value c' be called a vertex of C'. Since $\rho\rho' = cc'$, when $\eta = \eta'$, we have the

THEOREM: *If we establish a correspondence between any spherical curve C and any curve C' that is osculated by the quasi-osculating spheres of another curve, such that those points correspond for which $\eta = \eta'$, and if their respective vertices correspond, then their curvatures at corresponding points are inversely proportional.*

4. The equation $\rho = c\cos\eta$ holds for all curves lying on a sphere of radius c, provided $\rho = c$, when $\eta = 0$. To obtain an equation that holds

for *all* spherical curves and is independent of the point where $\eta = 0$, we must differentiate twice and eliminate c and η, giving

$$\frac{d^2\rho}{d\eta^2} + \rho = 0. \tag{1}$$

Conversely, every solution of (1) is of the form $\rho = c\cos(\eta + \varkappa)$; hence (1) is a characteristic property of spherical curves. For the classes J_2 and K_2, similarly, we must eliminate c and s, and for K eliminate c and θ. The results are easily obtained and are expressed by the

THEOREM: *Each of the following differential equations of the second order is a necessary and sufficient condition that a curve belong to the corresponding indicated class*

$$(J_1)\ \frac{d^2\rho}{d\eta^2} + \rho = 0, \qquad (J_2)\ \frac{d^2s}{d\sigma^2} = 0,$$

$$(K_2)\ \rho\frac{d^2\rho}{d\eta^2} - 2\frac{d\rho}{d\eta} - \rho^2 = 0, \qquad (K_2)\ \sigma(1+\sigma^2)\frac{d^2s}{d\sigma^2} + 2\frac{ds}{d\sigma} = 0,$$

$$(K)\ 3\rho\frac{d^2\rho}{d\theta^2} - 4\frac{d\rho}{d\theta} - 9\rho^2 = 0.$$

5. Let us now consider the curves that are common to two or more classes. In the first place it is obvious that any two of the classes I_1, J_1, K_1 (§ 2) have no curves in common; *they are mutually exclusive classes.* The same thing is true of the dual classes I_2, J_2, K_2. Moreover, I_1 and K are mutually exclusive. But with these exceptions any two of the classes have a two or three-parameter family of curves in common.

With respect to the first five classes I, I_1, I_2, J_1, J_2, the results are largely known. E. g., I, I_1, and I_2 have in common the two-parameter family of circular helices; and spherical curves of constant torsion (IJ_1) and of constant slope (I_2J_1) have been studied.* We shall therefore confine our attention principally to curves belonging to the new classes K_1, K_2, and K. In order to find the number of parameters on which a curve common to two classes depends, we have to express ρ and τ (or ρ and σ, or τ and σ) in terms of the arc-length s.

6. We begin with the family (I_2K_1), each curve of which is a cylindrical helix osculated by the quasi-osculating spheres of another curve, and satisfying the equations

$$\rho = c\sec\eta, \quad \sigma = a = -\cot\omega \qquad (c>0,\ a \neq 0), \tag{2}$$

* Cesàro, Vorlesungen über Natürliche Geometrie (1901), § 139, *d* and *h*.

where ω, by M. A., (22), § 10, is the constant angle the curve makes with a fixed direction. Its torsion evidently has a constant sign. By M. A., (19), § 7, we have

$$c\frac{d\varrho}{ds} = \pm\sigma\sqrt{\varrho^2 - c^2} = \pm a\sqrt{\varrho^2 - c^2},$$

$$\varrho = c\cosh\frac{as}{c}, \qquad \tau = \frac{c}{a}\cosh\frac{as}{c}, \tag{3}$$

where $s = 0$ at the vertex ($\varrho = c$, $\eta = 0$). Hence $(I_2 K_1)$ is a two-parameter family, depending on a and c. From (2_1) and (3_1) we have

$$s = \frac{c}{a}\log(\sec\eta + \tan\eta). \tag{4}$$

Let us now for the moment use the subscript 1, as in M. A., §§ 3–10, to distinguish this curve C_1 from the curve of centers C of its osculating spheres (a twisted parabola). Then (4), in view of M. A., (15), (§ 6), becomes

$$s_1 = \frac{c}{a}\log(\sec\eta_1 + \tan\eta_1) = -\frac{c}{a}\log(\sec\theta + \tan\theta),$$

which, in view of M. A., (5′), § 4, gives us

$$-as_1 = s - c\sec\theta\tan\theta$$

and also, since s and $c\sec\theta\tan\theta$ have like signs,

$$|as_1| = |s| - |c\sec\theta\tan\theta|,$$

and finally, by M. A., (6) and the figure, § 4,

$$|as_1| = |s| - NP.$$

This gives us the

THEOREM: *Let C_1 be a curve of constant slope a belonging to the family $(I_2 K_1)$ and C the curve of centers of its osculating spheres, and let their arc-lengths s_1, s at corresponding points P_1, P be measured from their respective vertices. Then the absolute value of as_1 is equal to the absolute value of s diminished by the projection of P_1P on the tangent to C at P.*

The *plane* curve defined by (3_1) was investigated by Cifarelli,* and in the special case where $a = \pm 1$, $\omega = \pm\pi/4$, it is the well-known "catenary of equal resistance".* Hence *every curve of type $(I_2 K_1)$ is a*

* See G. Loria, Spezielle Ebene Kurven (1902), § 236.

twisted curve of Cifarelli, and if the constant angle it makes with a fixed direction is $= \pi/4$, *it is a twisted "catenary of equal resistance".*

7. We proceed to find its *Cartesian* equations. Let the fixed direction with which it makes a constant angle ω be that of the z-axis. Then the direction-cosines of the tangent can be expressed in terms of a variable φ by the equations

$$\alpha = \sin\omega\cos\varphi, \quad \beta = \sin\omega\sin\varphi, \quad \gamma = \cos\omega. \tag{5}$$

Since $d\theta^2 = d\alpha^2 + d\beta^2 + d\gamma^2$, we can choose

$$d\theta = d\varphi \cdot \sin\omega, \quad \theta = \varphi\sin\omega, \tag{6}$$

where $\theta = 0$, when $\varphi = 0$. By (2_2), we have $d\eta/d\theta = \sigma = -\cot\omega$, whence $\eta = -\theta\cot\omega$, where $\eta = 0$, when $\theta = 0$. Hence, by (6_2), we have

$$\eta = -\varphi\cos\omega. \tag{7}$$

By means of (5), (6), and (7), the equations (43), of M. A., § 16, which are valid for *any* curve of class K_1, become in this particular case:

$$\begin{aligned} x &= c\sin^2\omega \int \cos\varphi \sec(\varphi\cos\omega)\, d\varphi, \\ y &= c\sin^2\omega \int \sin\varphi \sec(\varphi\cos\omega)\, d\varphi, \\ z &= c\sin\omega\cos\omega \int \sec(\varphi\cos\omega)\, d\varphi, \end{aligned} \tag{8}$$

where $c > 0$, $-\pi/2 < \omega < \pi/2$, $\omega \neq 0$. These are therefore the Cartesian equations of any curve of the family $(I_2 K_1)$. If $\cos\omega$ is a rational number (which must be positive and < 1), the integrals can be expressed in terms of elementary functions.

Putting $\eta = \pm\pi/4$, we obtain, for a curve of the family that is also a twisted "catenary of equal resistance", the equations

$$\begin{aligned} x &= \frac{c}{2}\int \cos\varphi \sec\frac{\varphi}{\sqrt{2}}\, d\varphi, \\ y &= \frac{c}{2}\int \sin\varphi \sec\frac{\varphi}{\sqrt{2}}\, d\varphi, \\ z &= \pm\frac{c}{2}\int \sec\frac{\varphi}{\sqrt{2}}\, d\varphi, \end{aligned} \tag{9}$$

which cannot be expressed in terms of elementary functions.

8. Coming next to the dual families $(J_2 K_1)$ and $(J_1 K_2)$, we shall find that each involves *three* parameters this time. By § 2 we see that any curve of $(J_2 K_1)$ has a fixed quasi-osculating sphere and variable osculating spheres that quasi-osculate another curve, and that it satisfies the equations

$$\varrho = c \sec \eta, \qquad s = a\sigma \qquad (c>0,\ a>0). \tag{10}$$

By M. A., (19), § 7, we have

$$\frac{d\varrho}{ds} = \pm \frac{\sigma}{c} \sqrt{\varrho^2 - c^2} = \pm \frac{s}{ac} \sqrt{\varrho^2 - c^2}.$$

Integrating, we have

$$\varrho = c \cosh \frac{s^2 - k}{2ac}, \qquad \tau = \frac{ac}{s} \cosh \frac{s^2 - k}{2ac},$$

where we cannot always choose $k = 0$, because the point from which s is measured is already determined by (10_2). There are two cases:

(a) If $k>0$, ϱ has its minimum value c at the two points $s = \pm\sqrt{k}$ and a relative maximum value, when $s = 0$.

(b) If $k \leqq 0$, ϱ has no maximum value, but has the minimum value $c \cosh (k/2ac)$, when $s = 0$.

In both cases the torsion vanishes at $s = 0$ and changes sign there; this is obviously the point at which the curve touches the fixed sphere (of radius a) quasi-osculating the curve.

9. Dually, a curve of type $(J_1 K_2)$ is a spherical curve quasi-osculated by the osculating spheres of another curve, and satisfying the equations

$$\varrho = a \cos \eta, \quad s = c(\sigma^{-1} - \sigma), \quad \text{where} \quad a>0,\ c>0. \tag{11}$$

Since

$$\varrho = \sigma\tau = \sigma \frac{ds}{d\eta} = -c(\sigma + \sigma^{-1}) \frac{d\sigma}{d\eta},$$

therefore

$$\cos\eta\, d\eta = -\frac{c}{a}(\sigma + \sigma^{-1})\, d\sigma,$$

$$\sin\eta = -\frac{c}{2a}[\sigma^2 + \log(k\sigma)^2], \tag{12}$$

where we can choose $k>0$. By means of (11) and (12) ϱ and σ are expressed implicitly in terms of s. We can also choose $\sigma>0$. At a vertex we have $\eta = n\pi$, $\varrho = \pm a$, and σ satisfies the equation

$$\sigma^2 + \log(k\sigma)^2 = 0,$$

which has just one finite positive solution. When $\eta = (2n+1)\pi/2$, then $\rho = \tau = 0$ and σ has either its maximum or its minimum value, which are the positive solutions of the respective equations

$$\sigma^2 + \log(k\sigma)^2 = \pm\frac{2a}{c},$$

and which correspond, by (11_2), to the respective minimum or maximum value of s.

Hence the curve consists of an infinite number of finite segments, all of the same length, connected with one another by cusps, at which both the curvature and the torsion become infinite, while σ is finite and $\neq 0$. The torsion is of constant sign on each segment and changes sign at each cusp. On each segment there is one interior point, the vertex, at which the osculating circle is a great circle of the sphere on which the curve lies. The entire curve is made up of congruent portions, each portion consisting of two adjacent segments.

10. Now consider the family (IK_1), which by § 2 consists of curves of constant torsion belonging to the class K_1.

For a curve C of this type we have

$$\rho = c\sec\eta, \qquad \tau = a,$$

and therefore

$$\eta = \frac{s}{a}, \qquad \theta = \frac{a}{c}\sin\frac{s}{a},$$

$$\rho = c\sec\frac{s}{a} = \pm\frac{ac}{\sqrt{a^2 - c^2\theta^2}}, \tag{13}$$

where $\eta = \theta = 0$, when $s = 0$. It is clear that C consists of an infinite number of segments, each of length πa, connected by points of inflection (at which the tangent is stationary). On each segment θ varies from $-a/c$ to a/c. The midpoint of each segment is a vertex, where $\rho = c$. The entire curve is made up of congruential repetitions of any two adjacent segments.

The *plane* curve C' defined by (13) is obtained by untwisting C, and its Cartesian equations are evidently

$$x = ac\int\frac{\cos\theta\, d\theta}{\sqrt{a^2 - c^2\theta^2}}, \qquad y = ac\int\frac{\sin\theta\, d\theta}{\sqrt{a^2 - c^2\theta^2}}.$$

11. The interesting self-dual curve of type $(J_1 J_2)$ will be discussed in a later paper. Another self-dual curve will now be considered, namely $C_1 (= C_2)$ of type $(K_1 K_2)$. By § 2, this remarkable curve, considered as C_2,

is quasi-osculated by the osculating spheres of a curve C_1'; and considered as C_1, it is osculated by the quasi-osculating spheres of a curve C_2'. Hence the curves of centers C and C' of its quasi-osculating and osculating spheres, respectively, are both twisted parabolas. We shall see that $C_1 (= C_2)$, and therefore also C_1', C_2', C, C', depends on three independent parameters. Obviously no two of these five curves can coincide, and any one of them, when chosen within its appropriate family, will uniquely determine the other four.

The intrinsic equations of a curve C_1 $(= C_2)$ are

$$\rho = c \sec \eta, \quad s = a(\sigma^{-1} - \sigma), \quad \text{where } c > 0, \ a > 0. \tag{14}$$

By M. A., (19), § 7, we have

$$\frac{d\rho}{d\sigma} = \pm \frac{\sigma}{c} \sqrt{\rho^2 - c^2} \, \frac{ds}{d\sigma} = \pm \frac{a}{c} (\sigma + \sigma^{-1}) \sqrt{\rho^2 - c^2},$$

and by integration

$$\rho = c \cosh \left[\frac{a}{2c} (\sigma^2 + 2 \log (b\sigma)) \right], \quad \text{where } b \neq 0. \tag{15}$$

By (14_2) and (15), ρ and σ are expressed implicitly as functions of s. The curve has one branch and two points at infinity. Its torsion is of constant sign, which is positive or negative according as $b > 0$ or < 0. When $\eta = 0$, then $\rho = c$ and $\sigma^2 + 2 \log (b\sigma) = 0$, which has one real solution in σ; hence the curve has one vertex. When $s = 0$, then $\sigma = \pm 1$ and

$$\rho = c \cosh \left[\frac{a}{2c} (1 + \log b^2) \right].$$

This point will coincide with the vertex, if and only if, $b = \pm 1/\sqrt{e}$.

If R and R' are the respective radii of the osculating and quasi-osculating spheres, then by M. A., (20), § 8, and (25_1), § 10, we have

$$R = \rho^2/c, \qquad R' = a(1 + \sigma^{-2}),$$

and by M. A., Theorem 4, § 8, and Theorem 7, § 12, these equations constitute a characteristic property of the curve. Hence we have the

THEOREM: *A curve will belong to the family* $(K_1 K_2)$, *if and only if,* (a) *the radii of its osculating spheres are variable and are proportional to* ρ^2, *and* (b) *the radii of its quasi-osculating spheres are proportional to* $1 + \sigma^{-2}$.

12. Let us now consider a twisted "catenary of equal resistance" (§ 6) belonging to the class (K_1), § 2. Its equations are

$$\rho = c \sec \eta = a \cosh (s/a), \text{ where } c > 0, \ a > 0.$$

By M. A., (19), § 7, we have

$$\tau = \pm \frac{a}{c} \coth \frac{s}{a} \cdot \sqrt{a^2 \cosh^2 \frac{s}{a} - c^2},$$

$$\theta = \sec^{-1} \frac{\varrho}{a}, \qquad \eta = \sec^{-1} \frac{\varrho}{c}.$$

There are three cases, as follows:

(a) $a = c$: here

$$\varrho = \pm \tau = c \cosh (s/c), \qquad \sigma = \pm 1,$$

so that the curve is also a cylindrical helix, and by §§ 6, 7, it is precisely the curve given by the parametric equations (9).

(b) $a > c$: when $s = 0$, and therefore $\varrho = a$, $\theta = 0$ and $\eta = \sec^{-1} \frac{a}{c}$, we get a singular point at which the osculator is stationary and the torsion vanishes and changes sign.

(c) $a < c$: when $\eta = 0$, and therefore $\varrho = c$, $\theta = \sec^{-1} \frac{c}{a}$, and $s = a \cosh^{-1} \frac{c}{a}$, we get a stationary point (cusp) at which the tangent is also stationary, but not the osculator, and the torsion becomes infinite and changes sign.

13. Finally, let us consider the curves that are common to the class (K) and some of the other classes listed in § 2. Of course these curves have no dual curves. We begin with a curve of type $(J_1 K)$, namely a spherical twisted parabola, whose equations are

$$\varrho = 2c \sec^3 \theta = a \cos \eta \qquad (a > 2c > 0),$$

where a is the radius of the sphere on which the curve lies, and θ is expressed in terms of s by M. A., (5′), § 4. Since $d\varrho/d\theta = \sigma \, d\varrho/d\eta$, we have

$$\sigma = \pm \frac{6c \sec^3 \theta \tan \theta}{\sqrt{a^2 - 4c^2 \sec^6 \theta}},$$

$$\tau = \pm \frac{\sqrt{a^2 - 4c^2 \sec^6 \theta}}{3 \tan \theta}.$$

At the singular point where $s = \theta = \sigma = 0$ and $\eta = \cos^{-1} (2c/a)$, the osculator is stationary and ϱ has its minimum value $2c$; while at the stationary point (cusp), where

$$\eta = \sigma^{-1} = 0 \quad \text{and} \quad \theta = \sec^{-1} (a/2c)^{1/3},$$

the tangent is also stationary (compare § 12, case c) and ϱ has its maximum value a. At both singular points the torsion changes sign. The curve evidently consists of an infinite number of finite congruent segments connected by cusps.

14. Let us next take the family (J_2K) of quasi-spherical twisted parabolas, given by the equations

$$\varrho = 2c\sec^3\theta, \qquad s = a\sigma \qquad (c>0,\ a>0),$$
$$s = c[\sec\theta\tan\theta + \log(\sec\theta+\tan\theta)] + b. \tag{16}$$

(J_2K) is a three-parameter family; for we cannot necessarily choose $b=0$, because the point P where $s=0$ is already fixed by (16_2). P is the singular point ($\sigma=0$) at which the curve touches the fixed quasi-osculating sphere. On the other hand the vertex ($\theta=0$, $s=b$, $\varrho=2c$, $\tau=2ac/b$) is a non-singular point unless $b=0$, in which case it coincides with P. The curve always has two points at infinity.

15. Consider the family (K_1K). A curve $C_1(=C)$ of this family is a twisted parabola osculated by the quasi-osculating spheres of a curve C_2. Hence their curve of centers (the planar evolute of C_1) is another twisted parabola C'. C_1 is itself the curve of centers of a family of spheres that osculate a curve C_1' and quasi-osculate a curve C_2'. Any one of these five curves uniquely determines the other four. If ϱ and R are the radii of the osculating circle and osculating sphere, respectively, of C_1, then $R=\varrho^2/a$; and conversely, by M. A., Theorem 4, § 8, *any twisted parabola for which R is proportional to ϱ^2 is a curve of type (K_1K).*

Any such curve is given by equations of the form

$$\varrho = c\sec\eta = 2a\sec^3\theta \qquad (c>0,\ a>0),$$

where θ is expressed in terms of s by M. A., (5′), § 4. Since $d\varrho/d\theta = \sigma\, d\varrho/d\eta$, we have

$$\sigma = \pm\frac{3c\tan\theta}{\sqrt{4a^2\sec^6\theta - c^2}}.$$

Hence (K_1K) is a two-parameter family. Exactly as in § 12 there are three cases:

(a) $2a = c$: here

$$\sigma = \pm 3/\sqrt{\sec^4\theta + \sec^2\theta + 1},$$

and the torsion is of constant sign.

(b) $2a>c$: $s=\theta=0$ is a singular point at which the torsion vanishes and changes sign.

(c) $2a<c$: $\eta=0$ is a cusp at which the torsion becomes infinite and changes sign.

16. Consider the family (K_2K). A curve $C_2(=C)$ of this family is a twisted parabola quasi-osculated by the osculating spheres of a curve C_1. Their curve of centers is another twisted parabola C'. C_2 is itself the curve of centers of a family of spheres that osculate a curve C_1' and quasi-osculate a curve C_2'.

The equations of a curve $C_2(=C)$ are

$$\rho = 2c\sec^3\theta, \quad s = a(\sigma^{-1}-\sigma), \quad \text{where } c>0, \quad a>0,$$
$$s = c[\sec\theta\tan\theta+\log(\sec\theta+\tan\theta)]+b,$$

so that (K_2K) is a three-parameter family, somewhat like (J_2K) treated in § 14. However, the point $(s=0,\ \sigma=\pm1)$ and the vertex $(\theta=0,\ s=b,\ \rho=2c)$ are here both non-singular points.

If R is the radius of the quasi-osculating sphere of C_2, then $R=a(1+\sigma^{-2})$; and conversely, by M. A., Theorem 7, § 12, *any twisted parabola for which R is proportional to $1+\sigma^{-2}$ is a curve of type (K_2K).*

17. The last, but not least, family that we shall study is (I_2K), in which every curve C is a twisted parabola of constant slope. Its intrinsic equations are of the form

$$\rho = 2c\sec^3\theta, \quad \sigma = a = -\cot\omega \qquad (c>0,\ a\neq0)$$

where $|\omega|$ is the constant angle it makes with a fixed direction.

The corresponding curve C_1, by M. A., (14), § 6, is also of constant slope $1/a$, and therefore of type (I_2K_1). Hence, by § 6, its intrinsic equations are

$$\rho_1 = c\sec\eta_1 = c\cosh\frac{s_1}{ac}, \quad \sigma_1 = \frac{1}{a} = \cot\omega_1.$$

The respective cylinders on which C and C_1 are helices obviously have parallel generators, and since $\omega+\omega_1=\pm\pi/2$, the respective angles $|\omega|$ and $|\omega_1|$ which the curves make with the generators are complimentary. It is clear that the three corresponding curves C, C_1, C_2 have no (finite) singular points, and their torsions have the same constant sign, that of a.

18. Let us now find the *Cartesian* equations of a curve C of type (I_2K). Choosing the fixed direction to be that of the z-axis, we proceed as in § 7 and express the direction-cosines α, β, γ of the tangent in terms of a variable φ by equations (5), which, by means of (6), become

$$\alpha = \sin\omega\cos(\theta\csc\omega), \quad \beta = \sin\omega\sin(\theta\csc\omega), \quad \gamma = \cos\omega.$$

Hence the equations M.A., (45), § 16, which hold for *any* curve of class (K), become in this particular case

$$
\begin{aligned}
x &= 2c \sin\omega \int \cos(\theta \csc\omega) \sec^3\theta \, d\theta, \\
y &= 2c \sin\omega \int \sin(\theta \csc\omega) \sec^3\theta \, d\theta, \\
z &= 2c \cos\omega \int \sec^3\theta \, d\theta,
\end{aligned} \tag{17}
$$

which are the required Cartesian equations. If $\csc\omega$ is a rational number (which cannot be $= 1$), the coördinates are expressible as elementary functions of θ.

19. *Example.* Since the simplest allowable rational value of $\csc\omega$ is 2, let us now put $\csc\omega = 2$, $\omega = \pi/6$, $\sigma = -\sqrt{3}$, and perform the integrations in (17) under the assumption that the vertex $(\theta = 0)$ is at the origin. We obtain

$$
\begin{aligned}
x &= \frac{c}{2}[-\sec\theta\tan\theta + 3\log(\sec\theta + \tan\theta)], \\
y &= 2c(\sec\theta - 1), \\
z &= \frac{c}{2}[\sqrt{3}\sec\theta\tan\theta + \sqrt{3}\log(\sec\theta + \tan\theta)].
\end{aligned}
$$

We reduce this to a simpler form by applying the transformation

$$
x' = \frac{1}{2}(-x + \sqrt{3}\,z), \qquad y' = y, \qquad z' = \frac{1}{2}(\sqrt{3}x + z),
$$

which amounts to reflecting the curve in the yz-plane (and so changing the sign of σ) and then rotating it about the y-axis through an angle of $-\pi/3$. The result, after dropping the primes, is the following:

$$
\begin{aligned}
x &= c\sec\theta\tan\theta, \\
y &= 2c(\sec\theta - 1), \\
z &= \sqrt{3}\,c\log(\sec\theta + \tan\theta),
\end{aligned} \tag{18}
$$

for which $\omega = -\pi/6$ and $\sigma = -\cot(-\pi/6) = \sqrt{3}$. Hence (18) represents a twisted parabola C of positive torsion that makes an angle of $\pi/6$ with the direction of the line $x - \sqrt{3}\,z = 0$, $y = 0$. At the origin (vertex) the principal normal is the y-axis and the tangent is the line $\sqrt{3}\,x - z = 0$, $y = 0$. The curve is symmetric with respect to the y-axis, and lies entirely in the two octants $(+++)$ and $(-+-)$, corresponding to $\theta > 0$ and

$\theta < 0$, respectively. Besides being a helix on an oblique cylinder, C also lies on the following cylinders:

$$x = \frac{c}{2} \sinh \frac{2z}{\sqrt{3}\, c},$$

$$y + 2c = 2c \cosh \frac{z}{\sqrt{3}\, c},$$

$$16 c^2 x^2 = y(y+4c)\,(y+2c)^2.$$

By M. A., (7), § 4, the spheres

$$(X-x)^2 + (Y-y)^2 + (Z-z)^2 = c^2 \sec^4 \theta,$$

where x, y, z satisfy (18), will osculate the corresponding curve C_1 and quasi-osculate the corresponding curve C_2.

We proceed to find the Cartesian equations of C_1 and C_2. First, since $\rho\alpha = dx/d\theta$, etc., and $\rho = 2c \sec^3 \theta$, we find by differentiating (18) that

$$(19) \qquad \alpha = 1 - \frac{1}{2} \cos^2 \theta, \quad \beta = \sin \theta \cos \theta, \quad \gamma = \frac{\sqrt{3}}{2} \cos^2 \theta.$$

Again, since $l = d\alpha/d\theta$, etc., where l, m, n are the direction-cosines of the principal normal, we have

$$(20) \qquad \begin{aligned} l &= \cos \theta \sin \theta, \\ m &= \cos^2 \theta - \sin^2 \theta, \\ n &= -\sqrt{3} \cos \theta \sin \theta. \end{aligned}$$

By means of (18), (19), (20), and M. A., (11), § 6, we now find the equations of C_1 to be

$$(21) \qquad \begin{aligned} x_1 &= \frac{3}{2} c \sin \theta, \\ y_1 &= c(3 \cos \theta - 2), \\ z_1 &= \sqrt{3}\, c \left[\log (\sec \theta + \tan \theta) - \frac{3}{2} \sin \theta \right]; \end{aligned}$$

and we have $\omega_1 = -\pi/3$, $\sigma_1 = 1/\sqrt{3}$, so that C_1 makes an angle of $\pi/3$ with the direction of the line $x - \sqrt{3}\, z = 0$, $y = 0$. At the vertex $(0, c, 0)$, where $\theta = 0$, the tangent is $x + \sqrt{3}\, z = 0$, $y = c$. C_1, like C, is symmetric with respect to the y-axis. It lies on the elliptic cylinder

$$\frac{x^2}{(3c/2)^2} + \frac{(y+2c)^2}{(3c)^2} = 1$$

and has the two asymptotes $x = \pm 3c/2$, $y = -2c$, which are generators of this cylinder.

By means of (18), (19), and M. A., (29), § 13, we find the equations of C_2 to be

$$
\begin{aligned}
x_2 &= -\frac{c}{2}\csc\theta, \\
y_2 &= c(\sec\theta - 2), \\
z_2 &= \sqrt{3}\,c\left[\log(\sec\theta+\tan\theta) - \frac{1}{2}\csc\theta\right],
\end{aligned}
\tag{22}
$$

which is also symmetric with respect to the y-axis. It lies on the quartic cylinder given by

$$
(y+2c)^2 = \frac{c^2x^2}{x^2-c^2/4}, \quad \text{or } 4x^2 = \frac{c^2(y+2c)^2}{(y+2c)^2-c^2},
$$

which has four distinct sheets. The right branch of C_2 lies on the sheet for which $x<-c/2$, $y>-c$, and the left branch on the sheet for which $x>c/2$, $y>-c$. The line $\sqrt{3}\,x - z = 0$, $y = -c$, is the common asymptote of the two branches.

The curves (18), (21) and (22) are probably as simple examples of curves belonging to the respective classes (K), (K_1) and (K_2) as can be found, even if those classes contain any algebraic curves, which is doubtful.

CORNELL UNIVERSITY.

ON THE GENERALIZATION OF THE ALGEBRA OF THE THEORY OF NUMBERS. THE INVARIANCY OF INFINITE SERIES.*

By Oliver E. Glenn.

Fermat's theorem and the theory of indices relate to the recurrence of a function $\xi = a\eta$ after e iterations, ξ and η being indeterminates and a a fixed integer prime to an integral modulus n. The author has elsewhere extended this view-point to the case in which the function iterated is ξ as determined by

$$(1) \qquad F(\xi) = \sum_{j=0}^{m} \sum_{k=0}^{m-j} a_{jk}\, \xi^{m-j-k}\, \eta^k \equiv 0 \pmod{n}, \quad (a_{jk} \text{ integral}).$$

In this paper is considered the modular periodicities which are determined by sets of forms in r variables. The indices of periods are found by means of a class of modular resultants which are in some instances invariants under cyclic transformations. We establish sufficient conditions for periods in any prescribed modular domain δ. This is necessary since if the modulus is composite, for example, a congruence is conditioned by assuming the existence of a solution in δ. The requirements of the analogy between equations and congruences in relation to dialytic elimination make necessary the creation of a new type of number domain (§ I) defined by irreducible congruences (mod. n). Otherwise the number field for roots would be imperfectly known and restricted. If the modulus is prime, galoisian imaginary roots are admitted and there is no restriction of generality in this particular.

The congruence,

$$F(\xi, \eta) = a\xi + b\eta + c \equiv 0 \pmod{n},$$

determines a period of different residues $(\alpha_1, \cdots, \alpha_s)$ such that the following are consistent for some value of s:

$$F(\alpha_2, \alpha_1) \equiv 0, \quad F(\alpha_3, \alpha_2) \equiv 0, \quad \cdots, \quad F(\alpha_s, \alpha_{s-1}) \equiv 0, \quad F(\alpha_1, \alpha_s) \equiv 0 \pmod{n}.$$

* Received January 17, 1928. Presented to the American Mathematical Society, April and October 1927. During the progress of this research the facilities of the Library of Congress, Washington, D. C., were utilized.

The eliminant of this system conforms to the lemma in § I of this article and is of the form,

$$R_s = [a^s - (-b)^s]\alpha_s + c[a^s - (-b)^s]/(a+b) \equiv 0 \pmod{n}.$$

When we consider F to be symbolically solved for ξ we may write $\xi \equiv \Delta(\eta) \pmod{n}$. Then $\alpha_2 \equiv \Delta(\alpha_1)$, $\alpha_3 \equiv \Delta(\Delta(\alpha_1)) \pmod{n}$, $(\equiv \Delta^2(\alpha_1))$. The least period of the function $\xi \equiv \Delta(\eta)$ is the least s $(= e)$ for which $R_s \equiv 0 \pmod{n}$, whether identically or as a congruence in α_s. We then have,

$$\Delta^{e+1}(\alpha_1) \equiv \Delta(\alpha_1) \pmod{n}.$$

Thus the operator Δ, as applied to the integer α_1, modulo n, generates a cyclic group of order e whose operations are, $(\alpha_1 = \alpha)$,

$$I(\alpha), \Delta(\alpha), \Delta^2(\alpha), \cdots, \Delta^{e-1}(\alpha), [I(\alpha) = \Delta^e(\alpha) = \alpha].$$

By the use of two congruences, $F(\xi, \eta) \equiv 0$ and $G(\xi, \eta) = p\xi + q\eta + r \equiv 0 \pmod{n}$ we can obtain an abstract group of non-cyclic abelian type. Let the operator obtained from G be Ω. The necessary and sufficient condition for the relation,

$$\Delta\Omega(\alpha) = \Omega\Delta(\alpha),$$

is the consistency of the system of congruences,

$$F(\beta, \alpha) \equiv 0, \quad G(\gamma, \beta) \equiv 0, \quad G(\delta, \alpha) \equiv 0, \quad F(\gamma, \delta) \equiv 0 \pmod{n}.$$

Hence,

$$C = (a+b)r - (p+q)c \equiv 0 \pmod{n}.$$

If the period of Ω is e' then Δ, Ω determine a group of order ee' whose generating relations are,

$$\Delta^e(\alpha) = I(\alpha), \qquad \Omega^{e'}(\alpha) = I(\alpha), \qquad \Delta\Omega(\alpha) = \Omega\Delta(\alpha).$$

For example, if $n = 10$, $\alpha = 3$, $(e = 4, e' = 2)$, the following congruences determine an abelian group of order 8 and type [2, 1]:

$$F(\xi, \eta) = 7\xi + 4\eta + 3 \equiv 0, \quad G(\xi, \eta) = \xi + 6\eta + 1 \equiv 0 \pmod{10}.$$

More general interpretations relating to groups are given in a later section.

The transcendental nature of periods is indicated by the fact that R_s is an exponential function of s. A problem of probable significance, not studied in this paper, relates to the analogy with logarithms. Denote the period of F by $P(a, b, c)$ and let $P(\alpha, \beta, \gamma) = y$. To what extent is $P(a+\alpha, b+\beta, c+\gamma)$ determinate in terms of x and y?

In the latter part of this paper cyclic transformations T on an infinite number of variables are considered. Looking upon the terms of an infinite sequence as the coefficients of forms f in these variables, we construct concomitants of f under T. These are forms in the infinitude of variables and, properly, concomitants of the given series. Applications to numerical cases are noted.

I. Modular elimination.

Suppose n to be a positive integer, either prime or composite and consider m linear homogeneous congruences in m variables, with integral coefficients and determinant D:

$$\begin{aligned} &a_1 t + b_1 u + \cdots + l_1 z \equiv 0, \\ &a_2 t + b_2 u + \cdots + l_2 z \equiv 0, \\ &\cdots\cdots\cdots\cdots\cdots \\ &a_m t + b_m u + \cdots + l_m z \equiv 0 \pmod{n}. \end{aligned} \tag{2}$$

Conditions for the consistency of the set are stated here for reference.

LEMMA. *If there exists a set of integers $t, \cdots, z$, one, at least, prime to n, which satisfy the congruences of the set simultaneously, it is necessary that $D \equiv 0 \pmod{n}$. Conversely if $D \equiv 0 \pmod{n}$, being of rank $m-1$, the system is consistent for solution in integers not all zero $\pmod{n}$.*

In proof let t be the variable which is assumed prime to n. If all the cofactors A_i of elements a_i in D should be zero then $D \equiv 0$. In the contrary case we may multiply the respective congruences (2) by $A_i\,(i = 1, \cdots, m)$, and add, with the result $Dt \equiv 0 \pmod{n}$. Hence $D \equiv 0 \pmod{n}$ is a necessary condition.

Conversely if the rank of D is $m-1$ the congruences may be arranged so L_m, the cofactor of l_m, is incongruent to zero. We can then solve the first $m-1$ congruences for the first $m-1$ variables $t, \cdots, y$, obtaining for each a relation of the form

$$L_m t \equiv -|l_1 b_2 c_3 \cdots k_{m-1}| z \pmod{n}. \tag{3}$$

If L_m contains factors of n not found in some of the determinants $|a_1 b_2 \cdots l_j \cdots k_{m-1}|$ the impossibility of solution which would be thus caused can be restituted by choosing z as dz' where z' is an integer and d a properly chosen divisor of n. All congruences (3) will then have solutions. Multiply the last congruence (2) by L_m and use (3). It will be satisfied identically. The lemma is proved.

We shall refer to D as the eliminant of Stieltjes.* Note that if it is

* *Œuvres Complètes*, p. 317. Frobenius, Journ. für Math., 88 (1879), p. 96–116. Scarpis, Periodico di Mat., 23 (1908), p. 49–61. Dickson, *History of Theory of Numbers*, 2 (1919), p. 93.

required in addition that a certain variable $t, \cdots, y$ should be prime to n the corresponding congruence (3) must have a solution prime to n.

It is necessary for what follows that we adapt to congruences a theory of elimination between several equations in one unknown due to Stuyvaert and Dines. The re-stated theory applies to congruences exactly as it does to equations if n is prime and galoisian imaginary roots are admitted. But in order that it may hold, with restrictions, in all cases a theory of imaginary roots of congruences with a composite modulus is given.

A congruence, with $a_0, \cdots, a_m$ integral,

$$F(x) = a_0 x^m + a_1 x^{m-1} + \cdots + a_m \equiv 0 \pmod{n}, \tag{4}$$

may have no integral root or a number $\nu >, =$ or $< m$, as the case may be. Note that $x^2 + 7x + 2 \equiv 0 \pmod{10}$ has four integral roots, 1, 2, 6, 7. Factorization is not unique. Thus,

$$2x^2 - x + 4 \equiv (x-2)(2x+3) \equiv (x-6)(2x+1) \pmod{10}.$$

Complex domains. For every prime or composite modulus n there exists a quadratic,

$$f(x) = x^2 - x + r,$$

which has no integral roots (mod. n).

In proof, if $n = 2$ or 3 the respective forms $x^2 - x + 1 = g(x)$, $x^2 - x + 2$, have the property. Let $n > 3$ and consider the form $g(x)$. If it has an integral root a it has two viz., a and $-a+1$. Then among the numbers,

$$g(0),\ g(1),\ \cdots,\ g(n-1), \tag{5}$$

there are at least two zeros and hence at least one residue α of the set $1, 2, \cdots, n-1$ is missing from the set (5). Then will

$$f(x) = g(x) - \alpha = x^2 - x + r \equiv 0 \pmod{n},$$

have no integral root, which was to be proved.

We postulate the existence of an imaginary number ι for which

$$\iota^2 \equiv \iota - r \pmod{n}. \tag{6}$$

If we use (6) as an algebraic reduction formula every polynomial $F(\iota)$ from (4) is reduced to $\alpha\iota + \beta$; α, β integral. There are n^2 reduced forms including the residues $0, \cdots, n-1$.

Definition. The $n^2 - n$ imaginaries $\alpha\iota + \beta$, $(\alpha \not\equiv 0 \pmod{n})$ and the n integers $0, \cdots, n-1$, together, are said to form a quadratic complex domain Δ, (mod. n).

Other domains can be constructed.

A congruence (4) with or without integral roots may have imaginary roots in Δ, in number $>$, $=$ or $<m$. Thus with $r=1$ the linear congruence $2x+6\equiv 0 \pmod{10}$ has four and only four roots in Δ, viz. 2, 7 and $5\iota+2$, $5\iota+7$. The following quadratics with no integral roots in the domain $(0, 1, \cdots, 9)$ have only the indicated roots in Δ, (mod. 10, r = 1):

$$\begin{aligned} &x^2+7x+3, \quad \iota+6, \quad -\iota+7, \quad \iota+1, \quad -\iota+2, \\ &x^2+6x+1, \quad 4\iota+5, \quad -4\iota+9, \\ &2x^2-3x+7, \quad 4\iota+5, \quad -4\iota+9. \end{aligned}$$

The sets of roots in Δ of every congruence (4) is unique and does not exceed n^2 roots for any value of m however large. The fundamental fact leading to solutions is that $F(u+v\iota)$, $(u, v$ integral) is reducible by $f(\iota)$ to

$$P(u, v)\iota+Q(u, v)\equiv 0 \pmod{n}, \quad (P, Q \text{ real}). \tag{7}$$

With n given a finite number of tests give all pairs (u, v) for which, simultaneously, $P(u, v)\equiv 0$, $Q(u, v)\equiv 0 \pmod{n}$.

Definition. Two congruences (4) whose respective sets of roots in Δ are identical will be said to be compatible in Δ, and otherwise incompatible.

The definition may be applied to equations. Two equations may be compatible or not, for instance, in the domain of quadratic surds.

Resultant determinants. Consider the equalities formed from polynomials,

$$\begin{aligned} f(x) &= a_0 x^m + a_1 x^{m-1} + \cdots + a_m = 0, \\ g(x) &= b_0 x^q + b_1 x^{q-1} + \cdots + b_q = 0, \end{aligned}$$

$(a_0, b_0 \neq 0)$, with coefficients in a domain R and roots in a domain δ. They may be equations on the one hand or congruences on the other. If they are congruences, modulo n, a prime, R being the field of residues of n, δ is a galois field* $GF(n^e)$, and if they are congruences with a composite modulus n, δ may be Δ. If f and g have a common root a in δ, the linear equalities,

$$\begin{aligned} a^{q-1} f(a) &= 0, \quad a^{q-2} f(a) = 0, \cdots, f(a) = 0, \\ a^{m-1} g(a) &= 0, \quad a^{m-2} g(a) = 0, \cdots, g(a) = 0, \end{aligned}$$

in a^{m-q-1}, a^{m-q-2}, $\cdots$, a, 1 are consistent and the dialytic eliminant vanishes necessarily.

* Dickson, *Linear Groups, with an Exposition of the Galois Field Theory* (1901), p. 1.

Conversely assume that this eliminant E vanishes. Then there exist $m+q$ parameters $k_1, \cdots, k_{m+q}$ such that the following system of equalities is consistent for values of k_i in R.

$$\begin{aligned}
&a_0\, k_1 + b_0\, k_{q+1} = 0,\\
&a_1\, k_1 + a_0\, k_2 + b_1\, k_{q+1} + b_0\, k_{q+2} = 0,\\
&a_2\, k_1 + a_1\, k_2 + a_0\, k_3 + b_2\, k_{q+1} + b_1\, k_{q+2} + b_0\, k_{q+3} = 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_m k_q + b_q\, k_{m+q} = 0.
\end{aligned}$$

We multiply these equalities, respectively, by $x^{m+q-1}, x^{m+q-2}, \cdots, x$, 1 and add and obtain

$$\varphi(x)\, f(x) + \psi(x)\, g(x) = 0, \tag{8}$$

where

$$\begin{aligned}
\varphi(x) &= k_1\, x^{q-1} + k_2\, x^{q-2} + \cdots + k_q,\\
\psi(x) &= k_{q+1}\, x^{m-1} + k_{q+2}\, x^{m-2} + \cdots + k_{m+q}.
\end{aligned}$$

There results the following

THEOREM: *If $f(x) = 0$ has roots in δ a set of sufficient conditions that f and g have a common root in δ is that E vanish and that the forms of the pair (f, ψ) be incompatible, f having more roots in δ than ψ.*

The second condition always holds for equations if δ is the analytic complex domain since then two equations of different orders are incompatible. The second condition always holds for congruences modulo n, a prime, δ being a galois field. The latter case includes a theorem* due to Sylvester namely the above theorem when $f(x)$ has m integral roots, n being prime. In the present state of knowledge if δ is Δ, the modulus being composite, the second condition is retained.

In proof of the theorem substitute for x in (8) a number a of δ which is a root of $f(x)$ and not of $\psi(x)$. Then $g(x)$ annuled.

If the modulus is composite, (8) being a congruence, $\psi(a)$ may contain a divisor d of n. Then a annuls $g(x)$ (mod. n/d). It is then often useful to employ $dg(x)$ instead of $g(x)$.

In relation to the resultants used in this paper to obtain sufficient conditions the incompatibility in Δ of (f, ψ) is assumed when n is composite and will be referred to as condition I. When the resultants are matrices an appropriate set of conditions I will be assumed.

Suppose n either prime or composite and that three congruences in x are satisfied by a common root in δ,

$$f = ax^2 + bx + c \equiv 0, \quad g = a'x^2 + b'x + c' \equiv 0,$$
$$h = px^3 + qx^2 + rx + s \equiv 0 \pmod{n}.$$

* Amer. Journ. of Math., 2 (1879), p. 360.

Then the five congruences $xf \equiv 0$, $f \equiv 0$, $xg \equiv 0$, $g \equiv 0$, $h \equiv 0 \pmod{n}$ are satisfied by a common value of x in δ and the eliminants of Stieltjes, which are the determinants of order four of the matrix Q, vanish (mod. n):

$$Q = \begin{Vmatrix} a & 0 & a' & 0 & p \\ b & a & b' & a' & q \\ c & b & c' & b' & r \\ 0 & c & 0 & c' & s \end{Vmatrix}.$$

We express this briefly by the formula $Q \equiv 0 \pmod{n}$. Conversely, if $Q \equiv 0 \pmod{n}$ there exists a set of residues k_i, $(i = 1, \cdots, 5)$ such that, identically in x,

$$(k_1 x + k_2) f + (k_3 x + k_4) g + k_5 h \equiv 0 \pmod{n}.$$

Hence, subject to the corresponding conditions I, if one form, as f, has a root in δ it is a common root of the three congruences, or at least if n is composite, of the congruences with some of the forms multiplied by a divisor of n.

The generalizations of the theory of the matrix Q in the algebraic realm have been carried out by Stuyvaert* and by Dines.† The necessary adaptations, subject to the above theorem, give, in effect, a theory of matrix elimination and conditions for common solutions of any number of congruences of any orders. The following is the principal proposition of the modified theory:

THEOREM *Let*

$$F_j = C_{j0} X^{m_j} + C_{j1} X^{m_j - 1} + \cdots + C_{jm_j} \equiv 0 \pmod{n}, \quad (j = 1, \cdots, r),$$

be a set of congruences, C_{jk} being integers and n a prime or composite integer. Let $M_N(F_1, \cdots, F_r)$ be the matrix,

$$(9) \quad \left\| \begin{array}{ccccccccccc} C_{10} & & & C_{20} & & & \cdots & \overbrace{C_{j0} & & & N - m_j} & \cdots \\ C_{11} & C_{10} \cdots & & C_{21} & C_{20} \cdots & & & C_{j1} & C_{j0} \cdots & & & \\ \vdots & C_{11} \cdots & & \vdots & C_{21} \cdots & & & \vdots & C_{j1} \cdots & & & \\ C_{1m_1} & C_{1m_1 - 1} \cdots & C_{10} & C_{2m_2} & \vdots & C_{20} \cdots & & C_{jm_j} & \vdots & C_{j0} \cdots & \\ & C_{1m_1} \cdots & \vdots & & C_{2m_2} \cdots & & & & C_{jm_j} \cdots & & \\ & & C_{1m_1} & & & C_{2m_2} \cdots & & & & C_{jm_j} \cdots & \end{array} \right\|$$

* *Cinq Études de Géométrie Analytique* (1908), p. 60. Amer. Journ. of Math., 37 (1915).

† Amer. Journ. of Math., 35 (1913).

where N is any integer satisfying the condition $N \geqq m_i + m_j$, $(i \neq j)$. It is a matrix of N rows and $\nu = \sum_{j=1}^{r} (N - m_j)$ columns, and will be said to vanish when all of its determinants of order N are congruent (mod. n) to zero. Then there are $r-1$ "essential determinants" of M_N whose vanishing, modulo n, is equivalent to the vanishing of the matrix and is a necessary and a sufficient condition for the existence of one or more common roots in δ of the system $F_j \equiv 0$ (mod. n).

II. The modular periods of a pair of ternary polynomials.

A theory of the periods, modulo n, of two non-homogeneous forms in three variables, to be generalized with reference to the number of variables later in this paper, is given first in greater detail than is feasible when the problem is made general. Assume,

$$F(\xi, \eta, \zeta) = \sum_{i=l}^{0} \sum_{j=0}^{l-i} (a_{ij0}\eta^{l-i-j} + a_{ij1}\eta^{l-i-j-1}\zeta + \cdots + a_{ijl-i-j}\zeta^{l-i-j})\xi^i,$$

(10)

$$G(\xi, \eta, \zeta) = \sum_{i=m}^{0} \sum_{j=0}^{m-i} (b_{ij0}\eta^{m-i-j} + b_{ij1}\eta^{m-i-j-1}\zeta + \cdots + b_{ijm-i-j}\zeta^{m-i-j})\xi^i,$$

in which the coefficients a_{ijk}, b_{ijk} are integers. Let (α_1, β_1) be assigned, respectively, to η, ζ, supposing as a result,

$$F(\alpha_2, \alpha_1, \beta_1) \equiv 0, \qquad G(\beta_2, \alpha_1, \beta_1) \equiv 0 \pmod{n}.$$

Here α_1, β_1 are any chosen residues of n and α_2, β_2 are solutions in the integral domain δ, assumed to exist, of the congruences of order m in ξ,

$$F(\xi, \alpha_1, \beta_1) \equiv 0, \qquad G(\xi, \alpha_1, \beta_1) \equiv 0 \pmod{n}. \tag{11}$$

Next assign (α_2, β_2), in this order, to (η, ζ), and assume the existence of a pair of residues (α_3, β_3) such that,

$$F(\alpha_3, \alpha_2, \beta_2) \equiv 0, \qquad G(\beta_3, \alpha_2, \beta_2) \equiv 0 \pmod{n}.$$

Assign (α_3, β_3) to (η, ζ) and continue the process. The maximum number of integral pairs (α_r, β_r) which are available is n^2 and is finite.

Postulate. The congruences (10) from which the sets $\sigma_r : (\alpha_r, \beta_r)$ are determined, or the methods by which these sets are selected from among the various solutions for any value of r, satisfy prescribed conditions such that the determinations σ_r are single-valued.

Thus at any stage of the process we are dealing with a single branch of solutions σ_r and, as the process of developing a branch is continued, there must be a repetition, as when

$$\alpha_{r+s} \equiv \alpha_r, \qquad \beta_{r+s} \equiv \beta_r \pmod{n}.$$

The solutions from σ_r to σ_{r+s} then form a period of s pairs of residues. The entire branch may represented graphically by a circle and a line drawn tangent to it. The sets $\sigma_1, \cdots, \sigma_r$ are placed at equal distances on the line with σ_r at the point of tangency and $\sigma_r, \sigma_{r+1}, \cdots, \sigma_{r+s} = \sigma_r$ upon the circumference of the circle.

Branches with unitary periods. Assume $\sigma_r = \sigma_{r+1}$ and use the simpler notation $\sigma_r = (\alpha, \beta)$. The necessary and sufficient conditions are

$$F(\alpha, \alpha, \beta) \equiv 0, \qquad G(\beta, \alpha, \beta) \equiv 0 \pmod{n},$$

or

$$(12)\quad \begin{aligned} F(\alpha, \alpha, \beta) &= \sum_{i=l}^{0} \sum_{j=0}^{l-i} (a_{0ji} + a_{1ji} + \cdots + a_{l-i-jji})\, \alpha^{l-i-j} \beta^i, \\ G(\beta, \alpha, \beta) &= \sum_{i=m}^{0} \sum_{j=0}^{m-i} (a_{0jm-i-j} + a_{1jm-i-j-1} + \cdots + a_{m-i-jj0})\, \alpha^i \beta^{n-i-j}. \end{aligned}$$

If we eliminate β we obtain a dialytic resultant which, when expanded, is a polynomial of order lm in α with coefficients which are rational and integral in the numbers a_{ijk}, viz.,

$$(13)\qquad C = C_0 \alpha^{lm} + C_1 \alpha^{lm-1} + \cdots + C_{lm} \equiv 0 \pmod{n}.$$

THEOREM: *A necessary condition for a period $w = 1$ upon a branch of the simultaneous functions $F(\xi, \eta, \zeta)$, $G(\xi, \eta, \zeta)$ is that C should be congruent to zero for some residue α of n. Sufficient conditions for a period 1 in δ consist in $C \equiv 0$ and, if n is composite, the condition I.*

The condition I and the fact that the residues which form a period may be imaginary were referred to obscurely in theorems 2, 3 of the author's paper in Annals of Mathematics, volume 28, p. 371.

The case $l = m = 1$. If $l = m = 1$ we may assume,

$$\begin{aligned} F(\xi, \eta, \zeta) &= a\xi + b\eta + c\zeta + d, \\ G(\xi, \eta, \zeta) &= p\xi + q\eta + r\zeta + s. \end{aligned}$$

Then,

$$\begin{aligned} F(\alpha, \alpha, \beta) &= c\beta + (a+b)\alpha + d \equiv 0 \pmod{n}, \\ G(\beta, \alpha, \beta) &= (p+r)\beta + q\alpha + s \equiv 0 \pmod{n}, \end{aligned}$$

$$C = D_1 \alpha + E_1 = \begin{vmatrix} c & a+b \\ p+r & q \end{vmatrix} \alpha + \begin{vmatrix} c & d \\ p+r & s \end{vmatrix} \equiv 0 \pmod{n}.$$

The period t, $l = m = 1$. Before the theory is further generalized let us treat the case of a period t, $l = m = 1$. The hypotheses are,

$$
\begin{aligned}
F(\alpha_2, \alpha_1, \beta_1) &= a\alpha_2 + b\alpha_1 + c\beta_1 + d \equiv 0,\\
G(\beta_2, \alpha_1, \beta_1) &= p\beta_2 + q\alpha_1 + r\beta_1 + s \equiv 0,\\
F(\alpha_3, \alpha_2, \beta_2) &= a\alpha_3 + b\alpha_2 + c\beta_2 + d \equiv 0,\\
G(\beta_3, \alpha_2, \beta_2) &= p\beta_3 + q\alpha_2 + r\beta_2 + s \equiv 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
F(\alpha_1, \alpha_t, \beta_t) &= a\alpha_1 + b\alpha_t + c\beta_t + d \equiv 0,\\
G(\beta_1, \alpha_t, \beta_t) &= p\beta_1 + q\alpha_t + r\beta_t + s \equiv 0 \pmod{n}.
\end{aligned}
$$

Since this linear system of $2t$ congruences involves $2t$ unknowns we may eliminate all but α_1 by the method of Stieltjes, whence we have the determinant of order $2t$,

$$
(14)\quad C_t = \begin{vmatrix}
b\alpha_1+d & c & a & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot\\
 & d & \cdot & b & c & a & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot\\
 & d & \cdot & \cdot & \cdot & b & c & a & \cdots & \cdot & \cdot & \cdot & \cdot\\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & b & c & a & \cdot\\
a\alpha_1+d & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & b & c\\
q\alpha_1+s & r & \cdot & p & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot\\
 & s & \cdot & q & r & \cdot & p & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot\\
 & s & \cdot & \cdot & \cdot & q & r & \cdot & p & \cdots & \cdot & \cdot & \cdot & \cdot\\
 & s & \cdot & \cdot & \cdot & \cdot & \cdot & q & r & \cdots & \cdot & \cdot & \cdot & \cdot\\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & q & r & \cdot & p\\
 & s & p & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & q & r
\end{vmatrix} \equiv 0 \pmod{n}.
$$

For brevity we write,

$$C_t = D_t\alpha_1 + E_t \equiv 0 \pmod{n},$$

where D_t, E_t are determinants whose form is evident from C_t, and C_1 is C. We conclude

THEOREM. *The least period of the set of two linear forms F, G is the subscript of the first resultant which can vanish, modulo n, in the series $C_1, \cdots, C_t, \cdots$.*

The formula C_2 is factorable,

$$C_2 = BC_1 = (cq - ap + bp + ar - br)\,C_1,$$

but the factorization cannot be generalized. If the linear pair F, G has no unitary period a necessary and sufficient condition for a binary period is the vanishing (mod. n) of the bilinear, biternary form B.

Analytically the two equations $F = G = 0$ (cf. (10)) represent the curve of intersection in 3-space of two surfaces, but, under the present circum-

stances, each pair (η, ζ) is associated with two different values of ξ. The theory is a numerical theory of the two abscissas drawn to two surfaces, respectively, from an arbitrary point (η, ζ) in the η, ζ plane. In the following section both real and imaginary abscissas are considered.

III. The periods of a set of $e-1$ non-homogeneous forms in e variables.

Given a set S of $e-1$ forms with integral coefficients, non-homogeneous in e variables $\xi_1, \cdots, \xi_e$,

$$S: F_i(\xi_1, \cdots, \xi_e), \qquad (i = 1, \cdots, e-1). \tag{15}$$

In F_i assign to $\xi_2, \cdots, \xi_e$, respectively, any $e-1$ residues, $(\alpha_1, \beta_1, \cdots, \varkappa_1)$, of n and suppose that $(\alpha_2, \beta_2, \cdots, \varkappa_2)$ are solutions in the domain δ of integral residues, assumed to exist, of the respective congruences,

$$F_1(\alpha_2, \alpha_1, \cdots, \varkappa_1) \equiv 0, \quad F_2(\beta_2, \alpha_1, \cdots, \varkappa_1) \equiv 0, \quad \cdots$$
$$\cdots, \quad F_{e-1}(\varkappa_2, \alpha_1, \cdots, \varkappa_1) \equiv 0 \pmod{n}.$$

Assign $(\alpha_2, \cdots, \varkappa_2)$ to $(\xi_2, \cdots, \xi_e)$ and assume,

$$F_1(\alpha_3, \alpha_2, \cdots, \varkappa_2) \equiv 0, \quad F_2(\beta_3, \alpha_2, \cdots, \varkappa_2) \equiv 0, \quad \cdots$$
$$\cdots, \quad F_{e-1}(\varkappa_3, \alpha_2, \cdots, \varkappa_2) \equiv 0 \pmod{n}.$$

Assign $(\alpha_3, \cdots, \varkappa_3)$ to $(\xi_2, \cdots, \xi_e)$ and continue the process. By postulate we consider only a single branch of solutions of the set S, which may be represented graphically by the figure described in II with $\sigma_r = (\alpha_r, \cdots, \varkappa_r)$. The maximum number of sets $(\alpha_j, \cdots, \varkappa_j)$ available for the solutions σ_r is n^{e-1}.

Let us assume that the t sets $(\sigma_1, \cdots, \sigma_t)$ form a characteristic circular period, that is, the least non-unitary period of S. The necessary and sufficient conditions are comprised in the simultaneity of the following $(e-1)t$ congruences:

$$F_1(\alpha_s, \alpha_{s-1}, \cdots, \varkappa_{s-1}) \equiv 0, \quad F_2(\beta_s, \alpha_{s-1}, \cdots, \varkappa_{s-1}) \equiv 0, \quad \cdots$$
$$\cdots, \quad F_{e-1}(\varkappa_s, \alpha_{s-1}, \cdots, \varkappa_{s-1}) \equiv 0 \pmod{n}, \ (s = 2, \cdots, t+1; \ \sigma_{t+1} = \sigma_1). \tag{16}$$

The number of parametric residues involved in this set is also $(e-1)t$. If we should arrange the forms of the set as algebroid quantics in $\alpha_t, \cdots, \iota_t, \alpha_j, \beta_j, \cdots, \varkappa_j$, $(j = 1, \cdots, t-1)$, whose coefficients are non-homogeneous polynomials in $\varkappa_t$, we could form Poisson's eliminant of the system obtaining as a result a rational integral polynomial in $\varkappa_t$ whose coefficients are rational and integral in the coefficients of the forms of the set. This resultant P, reduced modulo n, gives a necessary condition for consistency, viz.,

$$P(F_1, \cdots, F_{e-1}) \equiv Q(\varkappa_t) \equiv 0 \pmod{n}. \tag{17}$$

Since however, the method of elimination of Poisson is an algebraic process, usually described under the theory of symmetric functions, we proceed to develop a modified process which can be carried out practically by means of Stuyvaert's matrices.*

Consider sets $(\alpha_j, \cdots, \varkappa_j)$ in any modular domain δ. Let the system (16) be arranged as follows, with all forms assumed congruent to zero, modulo n:

$$
\begin{array}{l}
F_1(\alpha_2, \alpha_1, \cdots, \varkappa_1), \quad F_2(\beta_2, \alpha_1, \cdots, \varkappa_1), \quad \cdots, \quad F_{e-1}(\varkappa_2, \alpha_1, \cdots, \varkappa_1), \\
F_1(\alpha_3, \alpha_2, \cdots, \varkappa_2), \quad F_2(\beta_3, \alpha_2, \cdots, \varkappa_2), \quad \cdots, \quad F_{e-1}(\varkappa_3, \alpha_2, \cdots, \varkappa_2), \\
\cdots\cdots\cdots

eliminate γ_2, continuing until x_2 is eliminated. All ψ's will then have been utilized and the final set of determinants will be polynomials in $\alpha_3, \cdots, x_3$, $\alpha_t, \cdots, x_t$. These various steps are repeated until all sets $\alpha_j, \cdots, x_j$, $(j = 1, \cdots, t-1)$, have been eliminated. We then will have derived a set of polynomials $\alpha_t, \cdots, x_t$ alone,

$$P_1(\alpha_t, \cdots, x_t), \quad P_2(\alpha_t, \cdots, x_t), \cdots, P_\eta(\alpha_t, \cdots, x_t).$$

Their Stuyvaert matrix for the elimination of α_t gives the set,

$$Q_1(\beta_t, \cdots, x_t), \quad Q_2(\beta_t, \cdots, x_t), \cdots, Q_\zeta(\beta_t, \cdots, x_t).$$

Continuation gives the result of the final step in the form of a single resultant which is a polynomial in x_t, $Q(x_t) \equiv 0 \pmod{n}$. This polynomial, algebraically considered, cannot be other than Poisson's eliminant. Moreover, we can prove, for any modular domain δ,

THEOREM. *The existence in δ of a residue x_t to satisfy the congruence $Q(x_t) \equiv 0$ (mod. n) is a necessary condition for the existence of at least one whole system $(\alpha_j, \cdots, x_j)$, $(j = 1, \cdots, t)$ which satisfies the system* (18), *forming a period of t sets. A complete set of sufficient conditions consists of $Q(x_t) \equiv 0$, and, if n is composite, the set of conditions I.*

In fact the vanishing modulo n, of any one of the sets of determinants derived, such as (19), is thus necessary and sufficient for the existence, in δ, of the parameter eliminated at that stage of the method, this parameter satisfying the previous set of congruences from which it is eliminated. Thus, beginning with $Q(x_t) \equiv 0 \pmod{n}$, we can trace back to the original system (18), arriving there with a complete set $(\alpha_j, \cdots, x_j)$, $(j = 1, \cdots, t)$, whose existence to satisfy (18) is compelled by the vanishing of $Q(x_t)$. The system $F_1, \cdots, F_{e-1}$, then possesses a period of t real or imaginary sets $(\alpha_1, \cdots, x_1), \cdots, (\alpha_t, \cdots, x_t)$.

The case $e = 3$, $t = 2$: Formal development. To find the condition Q for a binary period (α', β'), (α, β) of the pair,

$$F_1 = a\xi^2 + b\xi\eta + c\eta^2 + d\xi\zeta + e\eta\zeta + f\zeta^2 + g\xi + h\eta + i\zeta + j,$$
$$F_2 = p\xi + q\eta + r\zeta + s,$$

we first eliminate α' by a Stuyvaert matrix of three columns and four rows, obtained from $F_1(\alpha, \alpha', \beta')$, $F_2(\beta, \alpha', \beta')$. Then β' is eliminated from two third order determinants of this matrix, $\varphi_1(\alpha, \beta, \beta')$, $\varphi_2(\alpha, \beta, \beta')$, the eliminant being a matrix M of three columns and four rows. Two determinants of M, of order three, $X_1(\alpha, \beta)$, $X_2(\alpha, \beta)$, give the resultant $Q(\beta)$, a determinant of the fourth order,

$$Q(\beta) = \begin{vmatrix} A & B\beta+C & D\beta^2+E\beta+F & 0 \\ 0 & A & B\beta+C & D\beta^2+E\beta+F \\ A' & B'\beta+C' & D'\beta^2+E'\beta+F' & 0 \\ 0 & A' & B'\beta+C' & D'\beta^2+E'\beta+F' \end{vmatrix},$$

in which

$$\begin{aligned}
A &= ap^2q^2+bpq^2r+cq^2r^2-dpq^3-eq^3r+fq^4,\\
B &= bpqr^2-bp^3q-2cp^2qr-dpq^2r+ep^2q^2+2cqr^3-2eq^2r^2+2fq^3r,\\
C &= bpqrs-bp^2qs+2cqr^2s-2cpqrs-dpq^2s+epq^2s-2eq^2rs\\
&\quad +2fq^3s+gp^2q^2+hpq^2r-ipq^3,\\
D &= cp^4-2cp^2r^2+cr^4+ep^2qr-eqr^3+fq^2r^2,\\
E &= 2cp^3s-2cpr^2s-2cp^2rs+2cr^3s+epqrs-2eqr^2s+ep^2qs\\
&\quad +2fq^2rs+hpqr^2-hp^3q-ipq^2r,\\
F &= cp^2s^2-2cprs^2+cr^2s^2+epqs^2-eqrs^2+hpqrs-hp^2qs-ipq^2s\\
&\quad +fq^2s^2+jp^2q^2,\\
A' &= aq^2r^2+bpq^2r+cp^2q^2,\\
B' &= 2aqr^3-2ap^2qr+bpqr^2-bp^3q+dpq^2r+ep^2q^2,\\
C' &= 2aqr^2s-2apqrs+bpqrs-bp^2qs+gpq^2r+hp^2q^2,\\
D' &= ap^4-2ap^2r^2+ar^4+dpqr^2-dp^3q+fp^2q^2,\\
E' &= 2ap^3s-2apr^2s-2ap^2rs+2ar^3s+dpqrs-dp^2qs+gpqr^2\\
&\quad -gp^3q+ip^2q^2,\\
F' &= ap^2s^2-2aprs^2+ar^2s^2+gpqrs-gp^2qs+jp^2q^2.
\end{aligned}$$

Numerical periods. The general theory has its arithmetical as well as its formal phase. Consider the numerical problem of finding the binary periods of the pair,

$$\begin{aligned}
F &= \xi^2+7\xi\eta+\eta^2+3\eta\zeta+2 \equiv 0,\\
G &= 9\xi+\eta+3\zeta+1 \equiv 0 \pmod{10}.
\end{aligned}$$

Using, for simplicity, $\alpha_1 = w$, $\beta_1 = x$, $\alpha_2 = y$, $\beta_2 = z$, we have,

$$\begin{aligned}
F(y, w, x) &= y^2+7yw+w^2+3wx+2 \equiv 0,\\
G(z, w, x) &= 9z+w+3x+1 \equiv 0 \pmod{10},\\
F(w, y, z) &= w^2+7wy+y^2+3yz+2 \equiv 0,\\
G(x, y, z) &= 9x+y+3z+1 \equiv 0 \pmod{10}.
\end{aligned}$$

The w eliminant of the first three congruences is the matrix,

$$\begin{Vmatrix} 1 & 3x+7y & y^2+2 \\ 1 & 7y & y^2+3yz+2 \\ 1 & 3x+9z+1 & 0 \\ 0 & 1 & 3x+9z+1 \end{Vmatrix}.$$

Two essential determinants, expanded, are

$$\varphi_1 = 9x^2+y^2+z^2+9xy+4xz+6x+3y+8z+3 \equiv 0,$$
$$\varphi_2 = y^2+z^2+9xy+7xz+7yz+3x+3y+8z+3 \equiv 0 \pmod{10}.$$

The x eliminant of $\varphi_1, \varphi_2, G(x, y, z)$, is,

$$\begin{Vmatrix} -1 & -y+4z+6 & y^2+z^2+3y+8z+3 \\ 0 & -y+7z+3 & y^2+z^2+7yz+3y+8z+3 \\ -1 & y+3z+1 & 0 \\ 0 & -1 & y+3z+1 \end{Vmatrix}.$$

Two determinants of this matrix give, when expanded,

$$\psi_1 = (z+5)y+2z^2+4z+6 \equiv 0,$$
$$\psi_2 = -y^2+5yz+4z^2+6y+4z+8 \equiv 0 \pmod{10},$$

and the dialytic y eliminant of ψ_1, ψ_2 is

$$6z^3+4z^2+6z+4 \equiv 0 \pmod{10}.$$

The real solutions for z are 1, 2, 3, 6, 7, 8, and, with z assigned, y is determined from ψ_1, x from $G(x, y, z)$ and w from $G(z, w, x)$. The following is the complete set of real periods obtained, four being binary,

$$[(w, x), (y, z)] = [(3,3), (3,3)], [(9,7), (3,1)],$$
$$[(4,2), (8,1)], [(8,8), (8,3)], [(4,7), (8,6)].$$

The algebraic eliminant of Poisson. If the forms in the system (18) are algebraic instead of modular and the matrices employed are algebraic the final result $Q(x_t)$ is the explicit form of Poisson's eliminant of this system. This is a new conclusion in the theory of elimination. The work of Stuyvaert and Dines was recent and the method of reaching the explicit form of the eliminant was not known in Poisson's day.

IV. The cyclical periods of a single form.

In connection with an e-ary form $F(\xi_1, \cdots, \xi_e)$ with integral coefficients, consider a necklace upon which are strung $e+\alpha-1$ numbers of a modular domain δ, $(\gamma_1, \cdots, \gamma_{e+\alpha-1})$. Let $\gamma_1, \cdots, \gamma_{e-1}$ be assigned to $\xi_1, \cdots, \xi_{e-1}$ respectively, and solve,

$$F(\gamma_1, \cdots, \gamma_{e-1}, \xi_e) \equiv 0 \pmod{n},$$

in δ for $\xi_e \equiv \gamma_e \pmod{n}$. Assign $\gamma_2, \cdots, \gamma_e$ to $\xi_1, \cdots, \xi_{e-1}$ and solve for $\xi_e \equiv \gamma_{e+1}$. Assign $\gamma_3, \cdots, \gamma_{e+1}$ to $\xi_1, \cdots, \xi_{e-1}$ and continue the process. Assume that the residues $\gamma_{e+\alpha}, \cdots, \gamma_{2e+\alpha-1}$, are respectively congruent to $\gamma_1, \cdots, \gamma_e$. Then the $(e+\alpha)$th step is identical with the first and we have $e+\alpha-1$ congruences in $e+\alpha-1$ parameters $\gamma_1, \cdots, \gamma_{e+\alpha-1}$ obtained by the indicated cyclical permutations of these parameters,

$$F(\gamma_1, \gamma_2, \cdots, \gamma_e) \equiv 0,$$
$$F(\gamma_2, \gamma_3, \cdots, \gamma_e, \gamma_{e+1}) \equiv 0, \quad \cdots, \quad F(\gamma_\alpha, \cdots, \gamma_{e+\alpha-1}) \equiv 0,$$
$$F(\gamma_{\alpha+1}, \cdots, \gamma_{e+\alpha-1}, \gamma_1) \equiv 0, \quad \cdots, \quad F(\gamma_{\alpha+e-1}, \cdots, \gamma_{e-1}) \equiv 0 \pmod{n}.$$

From these we can eliminate all parameters excepting γ_1 by forming the modular resultant (17) of Poisson,

$$P(F) = Q(\gamma_1) \equiv 0 \pmod{n}.$$

THEOREM. *For a fixed value of e the least period* $\geqq e$ *of* F, *in* δ, *is the least value of* $s = e+\alpha-1$ *for which* $Q(\gamma_1) \equiv 0$ *(mod. n).*

The case $m = 1$. According to the theory, if F is linear,

$$F = a_1\gamma_1 + a_2\gamma_2 + \cdots + a_e\gamma_e + b \equiv 0,$$

$$\begin{aligned}
&a_1\gamma_2 + a_2\gamma_3 + \cdots + a_e\gamma_{e+1} + b \equiv 0,\\
&\cdots\cdots\cdots

The determinant $D_{e+\alpha-1}$ is cyclical and therefore is, formally, linearly factorable in an irrational domain. Also if we add every column after the first to the first it is evident that,*

$$b\,D_{e+\alpha-1} = (a_1+a_2+\cdots+a_e)\,E_{e+\alpha-1}.$$

Suppose g to be a primitive $(e+\alpha-1)$st root of unity. Multiply the columns of $D_{e+\alpha-1}$, in order, by

$$1,\ g^t,\ g^{2t},\ \cdots,\ g^{(e+\alpha-2)t}, \tag{20}$$

where $t<e+\alpha-1$ is an integer, and there results, by addition of all new columns, respectively, to the first,

$$D_s = F_t(\lambda)\,D,\quad (s = e+\alpha-1),$$

in which D is the determinant obtained from D_s by replacing the elements of the first column of the latter by the numbers (20) respectively. Also,

$$F_t(\lambda) = \lambda_1+g^t\lambda_2+g^{2t}\lambda_3+\cdots+g^{(e+\alpha-2)t}\lambda_s,$$

where $\lambda_1,\cdots,\lambda_s$ are the elements of the first row of D_s. Since the principal diagonal of D_s is a_1^s it follows, in either the algebraic or modular case,

$$D_s = \prod_{t=0}^{s-1} F_t(\lambda). \tag{21}$$

The following are examples in expanded form; $e=3$.

It is worthy of note that they give an indefinite sequence of important symbolical identities when a_1, a_2, a_3, are replaced by the respective terms of the well-known identity, $(pq)r_x+(qr)p_x+(rp)q_x=0$:

$$\begin{aligned} D_3 &= a_1^3+a_2^3+a_3^3-3a_1a_2a_3,\\ D_4 &= a_1^4-a_2^4+a_3^4+4a_1a_2^2a_3-2a_1^2a_3^2,\\ D_5 &= a_1^5+a_2^5+a_3^5-5a_1a_2^3a_3+5a_1^2a_2a_3^2,\\ D_6 &= a_1^6-a_2^6+a_3^6+6a_1a_2^4a_3-9a_1^2a_2^2a_3^2+2a_1^3a_3^3. \end{aligned} \tag{22}$$

With $e=4$ the first of the sequence of expanded determinants is,

$$\begin{aligned} D_4 = {} & a_1^4-a_2^4+a_3^4-a_4^4-2a_1^2a_3^2+2a_2^2a_4^2-4a_1^2a_2a_4\\ & +4a_1a_2^2a_3-4a_2a_3^2a_4+4a_1a_3a_4^2. \end{aligned} \tag{23}$$

* Vandiver, Amer. Math. Monthly, 9 (1902), p. 96.

The methods relating to Theorem 1, II will give conditions for a period $<e$. In § VI D_s is considered as an invariant.

V. Modular periods and abstract groups.*

Let $\xi = Q(\eta)$, $\xi = R(\eta)$ be two functions obtained respectively from two modular polynomials (1), $F(\xi, \eta) \equiv 0$, $G(\xi, \eta) \equiv 0 \pmod{n}$, considered to be symbolically solved for ξ. Then, as was mentioned in the Introduction, under definite conditions, Q, R may be looked upon as operators of an abelian group of order equal to ee' the product of the periods of Q, R. We show that types of non-abelian groups may be represented in this way.

Consider the known abstract group of order $s^2 t$, where s, t are distinct primes, whose defining relations are,*

$$(24) \quad R^{s^2} = I, \quad Q^t = I, \quad R^{-1} Q R = Q^\beta, \quad \beta^{s^2} \equiv 1 \pmod{t}.$$

This type exists when $s = 2$, $t = 5$, $\beta = 2$. If Q, R are the modular operators determined from the respective linear congruences,

$$F(\xi, \eta) = a\xi + b\eta + c \equiv 0, \qquad G(\xi, \eta) = p\xi + q\eta + r \equiv 0 \pmod{n},$$

a necessary and sufficient condition that

$$QR(\alpha) = RQ^2(\alpha),$$

is evidently the consistency of the system,

$$G(\beta, \alpha) \equiv 0, \quad F(\gamma, \beta) \equiv 0, \quad F(\delta, \alpha) \equiv 0, \quad F(\varepsilon, \delta) \equiv 0, \quad G(\gamma, \varepsilon) \equiv 0 \pmod{n}.$$

The Stieltjes eliminant of this system is,

$$E = ap[bq(a+b)\alpha + ar(a+b) - cq(a-b) - acp] \equiv 0 \pmod{n}.$$

When $E \equiv 0 \pmod{n}$ and the indices of the respective periods of Q, R are 5, 4, while the periods themselves contain a common residue α, Q and R generate a group g of the type (24) and order 20. With $\alpha \equiv 1 \pmod{10}$ the following forms have the requisite properties and represent a group of this type and order,

$$F(\xi, \eta) = 3\xi + 2\eta + 7 \equiv 0, \qquad G(\xi, \eta) = \xi + 7\eta + 4 \equiv 0 \pmod{10}.$$

For these $E \equiv 0 \pmod{10}$ identically and respective modular periods are,

$$Q: (1, 7, 3, 9, 5), \qquad R: (1, 9, 3, 5).$$

* Hölder, Math. Annalen, 43 (1893), p. 335.

There results a graphical representation of the group g. We represent the operations of the cyclic group (Q) by rotations of the circle from one residue to another. The operations of (R) are routes of travel from 1 to other residues on the chords. The 20 elements of g are then routes of translation from 1 to other residues. They are all in the directions indicated by arrows, none being retrograde. The multiplication table of the group is as follows:

$$R^i(1),\quad QR^i(1),\quad Q^2R^i(1),\quad Q^3R^i(1),\quad Q^4R^i(1),\qquad (i=0,\cdots,3).$$

The route of travel represented by each operation is shown in the corresponding table below.

$$\begin{gathered}(1),\ (17),\ (173),\ (1739),\ (17395),\\ (19),\ (195),\ (19\overset{\frown}{51}),\ (19\overset{\frown}{517}),\ (19\overset{\frown}{5173}),\\ (193),\ (1939),\ (19395),\ (1939\overset{\frown}{51}),\ (1939\overset{\frown}{517}),\\ (1935),\ (193\overset{\frown}{51}),\ (193\overset{\frown}{517}),\ (193\overset{\frown}{5173}),\ (193\overset{\frown}{51739}).\end{gathered}$$

A general theory is obtained from (15) and the sets σ_r in § III. The set σ_2 is obtained by operating the system S upon σ_1 while $S^2\sigma_1 = S\sigma_2 = \sigma_3, \cdots, S^{t-1}\sigma_1 = \sigma_t$, $S^t\sigma_1 = \sigma_1$. Thus the operation S generates a cyclic group of order t. Necessary and sufficient conditions that operators S, T formed, respectively, from two such systems of congruences should generate an abelian group, are as follows:

(α) There must exist a set σ_i common to two periods such that, simultaneously,

$$(25)\qquad \begin{aligned} S\sigma_i &= (a,\cdots,k), \quad (=\sigma), \qquad T\sigma_i = (a',\cdots,k'), \quad (=\tau),\\ S\tau &= (a'',\cdots,k''), \quad (=v), \qquad T\sigma = v. \end{aligned}$$

(β) The analytic equivalent of the relation $S\tau = T\sigma$, in the form of Stieltjes' eliminant, must be congruent to zero modulo n.

If (α), (β) are satisfied and $T^s(\sigma_i) = I(\sigma_i) = \sigma_i$ we have a group of order st whose defining relations are as follows:

$$S^t(\sigma_i) = I(\sigma_i),\quad T^s(\sigma_i) = I(\sigma_i),\quad ST(\sigma_i) = TS(\sigma_i).$$

If the two systems are, for example, the following:

$$S:\begin{cases} a\xi + b\eta + c\zeta + d \equiv 0,\\ p\xi + q\eta + r\zeta + s \equiv 0, \end{cases}\qquad T:\begin{cases} a'\xi + b'\eta + c'\zeta + d' \equiv 0,\\ p'\xi + q'\eta + r'\zeta + s' \equiv 0 \end{cases}\quad (\text{mod. } n),$$

the eliminant whose vanishing (mod. n) gives $ST = TS$, is,

$$E = \begin{vmatrix} b & c & a & 0 & 0 & 0 & 0 & d \\ q & r & 0 & p & 0 & 0 & 0 & s \\ b' & c' & 0 & 0 & a' & 0 & 0 & d' \\ q' & r' & 0 & 0 & 0 & p' & 0 & s' \\ 0 & 0 & b' & c' & 0 & 0 & a' & d' \\ 0 & 0 & q' & r' & 0 & 0 & 0 & p'x+s' \\ 0 & 0 & 0 & 0 & b & c & a & d \\ 0 & 0 & 0 & 0 & q & r & 0 & px+s \end{vmatrix}.$$

VI. Invariants of cyclic transformations.

I have shown in a former paper* that the determinant D of the algebraic cyclic transformation on m variables,

$$C: \ (x) = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \\ \lambda_m & \lambda_1 & \cdots & \lambda_{m-1} \\ \cdot & \cdot & \cdots & \cdot \\ \lambda_2 & \lambda_3 & \cdots & \lambda_1 \end{bmatrix} (x'),$$

is factorable in the form,

$$D = \prod_{t=0}^{m-1} F_t(\lambda),$$

where,

$$F_t(\lambda) = \lambda_1 + g^t \lambda_2 + g^{2t} \lambda_3 + \cdots + g^{(m-1)t} \lambda_m, \quad (t = 0, \cdots, m-1),$$

and g is an arbitrary primitive mth root of unity. The complete system of universal covariants of C is the set $F_t(x)$, for which,*

$$F_t(x') = F_{(m-1)t}(\lambda)^{-1} F_t(x), \quad (t = 0, \cdots, m-1).$$

If the general algebraic form f of order p in m variables $x_1, \cdots, x_m$ is expanded in terms of the linear forms $F_t(x)$ as arguments, the $\binom{p+m-1}{p}$ coefficients of the expansion are linear forms, of the domain of g in the same number of coefficients of f. These forms compose a complete system of invariants of f under C in this domain.

Among the invariants of this system occur the sum $\sum a$ of the sequence of coefficients of f and the alternating series $\sum \pm a$ formed from this sequence. We find,

* Trans. Amer. Math. Soc., 18 (1917), p. 459.

$$(26)\qquad \begin{aligned} \sum a' &= (\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_m)^p \sum a, \\ \sum \pm a' &= (\lambda_1-\lambda_2+\lambda_3-\cdots-(-1)^m\lambda_m)^p \sum \pm a. \end{aligned}$$

Note that these two invariants are not functions of m.

This formal theory holds when C and f are modular, the modulus being n and g an integral primitive root.

Binary forms. Modular theory. There are however some restrictions on the modulus. If n is not a power of 2 and

$$f = a_0 x_1^p + a_1 x_1^{p-1} x_2 + \cdots + a_p x_2^p, \qquad (m=2),$$

the system of $p+1$ invariants is,

$$\varphi_i = \sum_{j=0}^{p} \sum_{t=0}^{i} (-1)^t \binom{p-j}{i-t}\binom{j}{t} a_j, \qquad (i=0,\cdots,p).$$

There is an incomplete relation between the cyclic invariants φ_i and the condition for a binary period $Q(x) \equiv 0 \pmod{n}$ of Theorem 1, III, expressed for one form in two variables, $(e=2)$. Here $Q(x)$ is the product of x^{m_3} and the dialytic eliminant I of f with

$$f_1 = a_p x_1^p + a_{p-1} x_1^{p-1} x_2 + \cdots + a_0 x_2^p.$$

In the determinant I if we add every column after the first to the first column the factor $\sum a$ is brought out. Then subtract the first row from all the other rows and we have a new determinant I_1 of order $2p-1$. Add the odd numbered columns of I_1 to the last column and the cyclic invariant $\sum \pm a$ appears as a factor of all elements of the new last column. Aside from these two factors I is not a cyclic invariant, $(p>2)$. Instead it is an invariant of such transformations of f as leave unaltered the property of possessing a binary period (mod. n).

The eliminant $Q(x)$ for $f = ax_1^2 + bx_1x_2 + cx_2^2$ is completely factorable in terms of cyclic invariants,

$$Q(x) = x^4(a+b+c)(a-b+c)(a-c)^2 = x^4 \varphi_0 \varphi_2 \varphi_1^2.$$

If n is an odd prime, $Q(x) \equiv 0$ and f has no unitary period we have either $a-c \equiv 0$ or $a-b+c \equiv 0 \pmod{n}$. Hence f is either symmetric or has a linear factor which is symmetric (mod. n), it being one of the types,

$$f = a_0(x_1^2+x_2^2) + a_1 x_1 x_2, \qquad f = (x_1+x_2)(a_0 x_1 + a_2 x_2).$$

The factors of $Q(x)$ for the cubic, $f = ax_1^3 + bx_1^2x_2 + cx_1x_2^2 + dx_2^3$, are as follows:

$$Q(x) = x^9(a+b+c+d)(a-b+c-d)[a(a-c)-d(d-b)]^3.$$

Cyclic periods. The determinant $D_{e+\alpha-1}$ of § IV is the logical generalization of Euler's binomials $a^s \pm b^s$.

THEOREM. *The determinant $D_{e+\alpha-1}$ is formally invariant, and is therefore an invariant modulo n, when*

$$F = a_1 x_1 + a_2 x_2 + \cdots + a_e x_e + a_{e+1} x_{e+1} + \cdots + a_m x_m,$$
$$(m = e + \alpha - 1,\ a_{e+1} = \cdots = a_m = 0),$$

is transformed by the cyclic transformation C in which $\lambda_{e+1} = \cdots = \lambda_m = 0$.

In proof let D be the determinant of (x). When F is transformed by (x) the coefficients of F' are, in order, the linear forms F with the x's replaced by the respective columns of D. Hence if $D'_{e+\alpha-1}$ is the corresponding transformed of $D_{e+\alpha-1}$ we have, $D'_{e+\alpha-1} = DD_{e+\alpha-1}$ by the formula for the product of two determinants.

The invariancy of infinite series. The sum and the alternating sum (26) remain invariant under cyclic transformations as the number of variables in the form is increased to infinity.

Given an infinite sequence,

$$a_0, a_1, \cdots, a_r, \cdots, \infty,$$

and the infinite cyclic transformation,

$$T: \begin{aligned} x &= \lambda_1 x' + \lambda_2 y' + \lambda_3 z' + \cdots, \infty, \\ y &= \lambda_1 y' + \lambda_2 z' + \lambda_3 v' + \cdots, \infty, \\ z &= \lambda_1 z' + \lambda_2 v' + \lambda_3 w' + \cdots, \infty, \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

We transform the quantic,

$$f_1 = a_0 x + a_1 y + a_2 z + \cdots, \infty,$$

by T. Assume an infinite convergent invariant series in the form

$$I = p a_0 + q a_1 + r a_2 + \cdots, \infty.$$

When an invariant relation for I is formed the method of undetermined coefficients gives the following unique determination, ξ being an arbitrary parameter:

$$I' = \sum_{i=0}^{\infty} a'_i \xi^i = \left(\sum_{j=1}^{\infty} \lambda_j \xi^{j-1}\right) \sum_{i=0}^{\infty} a_i \xi^i = \Delta I.$$

The formulae (26) are special cases if $m \doteq \infty$. The conditions for absolute convergency are assumed,

$$(27) \qquad \operatorname*{Lim}_{n \doteq \infty} \lambda_{n+1} \xi / \lambda_n < 1, \quad \operatorname*{Lim}_{n \doteq \infty} a_{n+1} \xi / a_n < 1.$$

THEOREM. *The sum of the series I and its property of convergence are invariants of f_1 under T. A theorem on the product of two power series in ξ is a theorem on the transformation of one (I) by the infinite cyclic transformation determined by the coefficients of the other.*

Consider next a quadratic form in an infinite number of variables,

$$f_2 = a_0 x^2 + a_1 y^2 + a_2 z^2 + \cdots + a_0' xy + a_1' xz + \cdots + a_0'' yz + a_1'' yt + \cdots, \infty.$$

It is easy to show, by partial differentiation, that if the set $\zeta_1, \zeta_2, \cdots, \infty$, is cogredient, under T, to the set $x, y, \cdots, \infty$, then the result of operating with

$$\Omega = \zeta_1 \frac{\partial}{\partial x} + \zeta_2 \frac{\partial}{\partial y} + \zeta_3 \frac{\partial}{\partial z} + \cdots, \infty,$$

upon any cyclic covariant of f_2 is a cyclic concomitant. In fact the polar operator Ω is the only fundamental invariantive algebraic process which survives when the number of variables becomes infinite.

We obtain a linear covariant,

$$F = \Omega f_2 = A_1 x + A_2 y + A_3 z + \cdots, \infty,$$

where

$$A_r = 2a_{r-1}\zeta_r + a_{r-2}'\zeta_1 + a_{r-3}''\zeta_2 + \cdots + a_0^{(r-1)}\zeta_{r-1} + a_0^{(r)}\zeta_{r+1} + a_1^{(r)}\zeta_{r+2} + a_2^{(r)}\zeta_{r+3} + \cdots, \infty, \quad (r = 1, 2, \cdots, \infty).$$

When F replaces f_1 its linear invariant constructed on the model of I is the concomitant,

$$J(\zeta_1, \zeta_2, \cdots) = \sum_{r=1}^{\infty} (a_{r-2}' + a_{r-3}''\xi + \cdots + a_0^{(r-1)}\xi^{r-2} + 2a_{r-1}\xi^{r-1} + a_0^{(r)}\xi^r + a_1^{(r)}\xi^{r+1} + a_2^{(r)}\xi^{r+2} + \cdots, \infty)\zeta_r.$$

A cyclic covariant of f_2 of the second degree in the $a_i^{(r)}$ is the polar,

$$\Omega f_2 J(x, y, \cdots, \infty).$$

The processes can be further extended.

Consider a form of order n in an infinite number of variables,

$$f_n = a_0 x^n + a_1 y^n + \cdots + a_0^{(1)} x^{n-1} y + \cdots, \infty.$$

An invariant of first degree of f_n, which contains n arbitrary parameters may be determined by using a symbolical notation for f_n:

$$f_n = (\alpha_1 x + \alpha_2 y + \alpha_3 z + \cdots, \infty)^n.$$

Write

$$g(\alpha, \xi_i) = \alpha_1 + \alpha_2 \xi_i + \alpha_3 \xi_i^2 + \cdots, \infty,$$

and the desired invariant is

$$G(\xi_1, \xi_2, \cdots, \xi_n) = \prod_{i=1}^{n} g(\alpha, \xi_i).$$

The invariant relation is

$$G' = H(\lambda)\, G,$$

where,

$$H(\lambda) = \prod_{i=1}^{n} (\lambda_1 + \lambda_2 \xi_i + \lambda_3 \xi_i^2 + \cdots, \infty).$$

The actual form of G, when $n = 2$, is as follows, $(\xi_1 = \xi,\ \xi_2 = \eta)$,

$$G(\xi, \eta) = a_0 + \frac{1}{2} a_0'(\xi + \eta) + a_1 \xi\eta + \frac{1}{2} a_1'(\xi^2 + \eta^2)$$

$$+ \frac{1}{2} a_2'(\xi^3 + \eta^3) + \frac{1}{2} a_0''(\xi^2\eta + \xi\eta^2) + \cdots, \infty.$$

Among numerous numerical illustrations of these invariantive series we mention the geometric progression whose first term is unity and whose ratio is $r < 1$. The invariant series I is $1/(1 - \xi r)$. Also if all

$$(a_0^{(r)}, a_1^{(r)}, a_2^{(r)}, \cdots, \infty), \qquad (a_j^{(0)} = a_j),$$

in f_2 are the series,

$$m, m^2, m^3, \cdots, \infty, \qquad (m\xi < 1),$$

the covariant J is

$$J = \sum_{r=1}^{\infty} [m(m^{r-1} - \xi^{r-1})/(m - \xi) + 2\, m^r \xi^{r-1} + m \xi^r/(1 - m\xi)]\, \zeta_r.$$

THE UNIVERSITY OF PENNSYLVANIA.

INVOLUTIONS THAT BELONG TO A LINEAR CLASS.

By E. T. Browne.*

1. **Introduction.** An involution is a projective correspondence of period two so that the square of the corresponding matrix is a scalar, and conversely. Thus if† $A \neq cI$ is such that $A^2 = kI (k \neq 0)$, then A is the matrix of an involution. Since in projective correspondences a scalar multiplier is immaterial, we shall consider matrices A of order n which possess the property

$$A^2 = I, \qquad (A \neq \pm I). \tag{1}$$

It is well known‡ that a matrix A of period two has, say, $p (1 < p < n)$ characteristic roots at $+1$ and $n-p$ characteristic roots at -1, and that there exists a non-singular matrix S such that

$$S^{-1} A S = \begin{pmatrix} I_p, & 0 \\ 0, & -I_{n-p} \end{pmatrix}$$

where $I_p (-I_{n-p})$ stands for a $p (n-p)$ square matrix with $+1 (-1)$ in the main diagonal and 0 elsewhere. Further, corresponding to the characteristic root $+1$ there is a $p-1$ space S_A, the coördinates of each of whose points are unaltered by A, and corresponding to the characteristic root -1 an $n-p-1$ space T_A, the coördinates of each of whose points are merely changed in sign by A.

Since the coördinates of a fixed point of the transposed matrix A' are precisely the coördinates of a fixed plane of A, and since A and A' are similar, corresponding to the characteristic root $+1 (-1)$ there is under A a $p-1 (n-p-1)$ space $S'_A (T'_A)$ of fixed planes.

A point $(x_1, \cdots, x_n)$ is conveniently represented by a rectangular matrix X with one column. Using this notation, $AX = X$ for a point X of S_A and $AX = -X$ for a point X of T_A. Similarly, if a plane be represented by a matrix U with one row, $UA = U$ or $UA = -U$ according as U lies in S'_A or T'_A.

Every point of $S_A (T_A)$ lies on every plane of $T'_A (S'_A)$. For if X lies in $S_A (T_A)$ and U lies in $T'_A (S'_A)$, then $AX = X$, $UA = -U (AX = -X$, $UA = U)$, so that

$$-UX = UAAX = UA^2X = UX,$$

whence $UX = 0$.

* Received September 14, 1927, in revised form February 10, 1928.

† Here I is the *unit matrix* or the matrix of the *identical transformation*.

‡ Hilton, *Homogeneous Linear Substitutions*, Oxford, 1914, p. 32.

2. Involutions that belong to a linear class. *Definition.* If A and B are two square matrices of order n both of which possess the property (1) and if for every θ

$$C = (1-\theta)A + \theta B \tag{2}$$

also possesses the property (1), we shall say that A and B determine a *linear class* to which C belongs.

On forming the product C^2 and imposing the condition $C^2 = I$ for every θ, we have at once the theorem.

THEOREM I. *Two involutions A and B determine a linear class if, and only if,*

$$AB + BA = 2I. \tag{3}$$

Writing A' for the transpose of A and taking transposes of both members in (3) we have the immediate corollary.

Corollary 1. If A and B determine a linear class so also do A' and B' determine a linear class called the *transpose class* of that determined by A and B.

Corollary 2. If A and B determine a linear class and P is any non-singular matrix, then $P^{-1}AP = A_1$ and $P^{-1}BP = B_1$ determine a linear class.

Taking $\theta = 1/2$ in (2) the matrix $C = (A+B)/2$ is non-singular, being its own inverse. Since $AC = CB$, i. e., $C^{-1}AC = B$, A and B are similar. Hence corresponding to the characteristic roots $+1$, -1 of B fixed elements under B include $p-1$ spaces S_B of points and S'_B of planes, and $n-p-1$ spaces T_B of points and T'_B of planes. These spaces stand in the same relation to each other as the corresponding spaces of A (§ 1). Thus, A and B are of the same type geometrically. For, example, in space a line reflection and a point-plane reflection cannot determine a linear class.

3. Common fixed elements of A and B. The spaces S_A and T_B have no point X in common. For if $AX = X$, $BX = -X$, we have on adding $(A+B)X = 0$, which is impossible if $X \neq 0$, since by (3) $A+B$ is non-singular. Similarly, T_A and S_B have no point in common. Hence, a common fixed point of A and B is a common point of S_A and S_B or of T_A and T_B, and therefore satisfies the equations

$$AX = BX = \pm X. \tag{4}$$

Denote by M the matrix $A-B$. In view of (3) M evidently satisfies the conditions

$$AM = -MA, \tag{5}$$

$$M^2 = 0. \tag{6}$$

From (4) if X is any common fixed point of A and B then $MX = 0$, so that X belongs to the null space $\mathfrak{M}$ of M. Conversely, any common point of $S_A(T_A)$ and $\mathfrak{M}$ is a point of $S_B(T_B)$ and is therefore a common fixed point of A and B.

Suppose that S_A and $\mathfrak{M}$ have exactly σ linearly independent points $X_1, \cdots, X_\sigma$ $(0 \leqq \sigma \leqq p)$ in common. We may then choose as a basis of S_A

$$S_A: \quad (X_1, \cdots, X_\sigma, X_{\sigma+1}, \cdots, X_p),$$

where $MX_i = 0$ $(i = 1, \cdots, \sigma)$; $MX_{\sigma+i} = Y_i \neq 0$ $(i = 1, \cdots, p-\sigma)$. From the latter condition by (6) $MY_i = M^2X_{\sigma+i} = 0$, so that the Y_i are points of $\mathfrak{M}$. Also, by (5) $AY_i = AMX_{\sigma+i} = -MAX_{\sigma+i} = -MX_{\sigma+i} = -Y_i$, so that the Y_i are linearly independent points of T_A also, and are therefore common points of T_A and T_B.

Similarly, if T_A and $\mathfrak{M}$ have exactly τ $(0 \leqq \tau \leqq n-p)$ points $Y_1, \cdots, Y_\tau$ in common, where evidently $\tau \geqq p-\sigma$, we may choose as a basis of T_A

$$T_A: \quad (Y_1, \cdots, Y_\tau, Y_{\tau+1}, \cdots, Y_q), \qquad (q = n-p).$$

As before $MY_i = 0$ $(i = 1, \cdots, \tau)$; $MY_{\tau+i} = Z_i$ $(i = 1, \cdots, q-\tau)$, where the Z_i are linearly independent common points of S_A and $\mathfrak{M}$ and may therefore be chosen as X's. Since evidently $\sigma \geqq q-\tau$, we may state the following theorem.

THEOREM II. *If A and B determine a linear class and if A has a characteristic root of multiplicity p at $+1$, then A and B have in common at least as many points as the greater of the two numbers p and $n-p$.*

4. Number of common fixed points of A and B. Let us choose as a basis of our $n-1$ space the set of points:

$$\Sigma: \quad (X_1, \cdots, X_\sigma, X_{\sigma+1}, \cdots, X_p, Y_1, \cdots, Y_\tau, Y_{\tau+1}, \cdots, Y_q), \tag{7}$$

where the X's and Y's are those of section 3. Under A these points are transformed into the points:

$$\Sigma_{(A)}: (X_1, \cdots, X_\sigma, X_{\sigma+1}, \cdots, X_p, -Y_1, \cdots, -Y_\tau, -Y_{\tau+1}, \cdots, -Y_q). \tag{8}$$

Under M these points Σ are transformed into:

$$\Sigma_{(M)}: (0, \cdots, 0, \quad Y_1, \cdots, Y_{p-\sigma}, \quad 0, \cdots, 0, \quad X_1, \cdots, X_{q-\tau}). \tag{9}$$

From (9) the rank of M is evidently $r = n-(\sigma+\tau)$. We therefore have the theorem.

THEOREM III. *If A and B determine a linear class and if r is the rank of $M = A - B$, A and B have exactly $n - r$ fixed points in common.*

From the conditions $\sigma+\tau \geqq p; \geqq n-p$ we have the immediate corollary.

Corollary. The rank r of M cannot exceed the lesser of the two numbers $p, n-p$.

5. **Distribution of the common fixed points of A and B.** Consider the auxiliarly matrices E, F defined as follows:

$$E = (A-I)(B+I) = M - AM; \qquad F = (B+I)(A-I) = M - MA;$$

which in view of (3) and (5) may be written in the equivalent forms:

$$E = -(B-I)(A+I) = M + MA; \quad F = -(A+I)(B-I) = M + AM.$$

Using (9) it is clear that the basis Σ is transformed under AM as follows:

$$(10) \quad \Sigma_{(AM)}(0, \cdots, 0, -Y_1, \cdots, -Y_{p-\sigma}, 0, \cdots, 0, X_1, \cdots, X_{q-\tau}).$$

Hence, since $E = M - AM$, from (9) and (10) $\Sigma_{(E)}$ is

$$(11) \quad \Sigma_{(E)}(0, \cdots, 0, 2Y_1, \cdots, 2Y_{p-\sigma}, 0, \cdots, 0, 0, \cdots, 0).$$

Similarly, since $F = M + AM$, $\Sigma_{(F)}$ is

$$(12) \quad \Sigma_{(F)}(0, \cdots, 0, 0, \cdots, 0, 0, \cdots, 0, 2X_1, \cdots, 2X_{q-\tau}).$$

Hence the ranks of E and F are $r' = p - \sigma$ and $r'' = q - \tau = n - p - \tau$.

THEOREM IV. *If r', r'' are the ranks of the auxiliary matrices $E = (A-I)(B+I)$, $F = (B+I)(A-I)$, respectively, the spaces S_A and S_B have exactly $\sigma = p - r'$ points in common, while the spaces T_A and T_B have exactly $\tau = n - p - r''$ points in common.*

There follow at once the following corollaries.

Corollary 1. A necessary and sufficient condition that $S_A(T_A)$ coincide with $S_B(T_B)$ is that $r'(r'') = 0$; i. e., $E(F) = 0$.

Corollary 2. If S_A does not coincide with S_B, T_A and T_B have at least one point in common.

From Theorems III and IV we have also the corollary.

Corollary 3. If r, r' and r'' are the ranks of M, E and F, respectively, then

$$(13) \quad r' + r'' = r.$$

6. Certain common fixed points of A and B.

THEOREM V. *If A and B determine a linear class the columns of $E(F)$ are the coördinates of $r'(r'')$ linearly independent common points of T_A and T_B (S_A and S_B) which lie not only on every common fixed plane of A and B but also on every plane of each of the spaces S'_A, S'_B (T'_A, T'_B).*

For let X be a point whose coördinates are a column of E. Since $(A+I)E = (B+I)E = 0$, X is a common point of T_A and T_B.

If U is a common plane of the spaces T'_A, T'_B so that $UA = UB = -U$, then $UM = 0$ and $UE = U(M - MA) = 0$.

If U is a plane of either of the spaces S'_A, S'_B so that either $U(A-I) = 0$ or $U(B-I) = 0$, then obviously, $UE = 0$.

The part of the theorem which relates to the columns of F may be proved in a similar manner.

7. Common fixed planes of A and B. The coördinates of a fixed point of A' are precisely the coördinates of a fixed plane of A. Since A' and B' determine a linear class, the common fixed points of A' and B' are precisely the common fixed planes of A and B. For this transpose class $M' = A' - B'$ plays the role of M. However, $E' = M' - M'A'$ plays the role of F, and F' plays the role of E. Hence, if in the theorems stated above concerning common fixed points of A and B we replace the word *point* by *plane*, the spaces S_A, S_B, by S'_A, S'_B, etc., replace E by F', F by E', and interchange r' and r'', we obtain corresponding theorems concerning the common fixed planes of A and B.

8. Fixed elements of $C = (1-\theta)A + \theta B = A - \theta M$. Choosing as the basis of our $n-1$ space the set of points Σ in (7) and using (8) and (9), we find that under $C = A - \theta M$ Σ is transformed into the set

$$\Sigma_{(C)}\ (X_1, \cdots, X_\sigma,\ X_{\sigma+1} - \theta Y_1, \cdots, X_p - \theta Y_{p-\sigma},\ -Y_1, \cdots$$
$$\cdots, -Y_\tau,\ -Y_{\tau+1} - \theta X_1, \cdots, -Y_q - \theta X_{q-\tau}).$$

The points of the spaces S_C and T_C are therefore:

$$S_C\colon\ X_1, \cdots, X_\sigma;\ \text{and}\ \ X_{\sigma+i} - \frac{\theta}{2} Y_i, \qquad (i = 1, \cdots, p-\sigma),$$

$$T_C\colon\ Y_1, \cdots, Y_\tau;\ \text{and}\ \ Y_{\tau+i} + \frac{\theta}{2} X_i, \qquad (i = 1, \cdots, q-\tau).$$

Since $X_{\sigma+i}$ are points of S_A (but not of S_B), Y_i are common points of T_A and T_B, and $X_{\sigma+i} - \frac{1}{2} Y_i$ are points of S_B* (but not of S_A), we may state the following theorem.

* Since $B = A - M$.

THEOREM VI. *Every fixed point P of S_C (T_C) is either a common point of S_A and S_B (T_A and T_B) or lies on a line joining a point Q of S_A (T_A) to a point S of S_B (T_B) and passing through a common point R of T_A and T_B (S_A and S_B). Every point on this line is a fixed point of some involution of the class, every point except R arising from the characteristic root $+1$ (-1), and the cross ratio of any four of these points (except R) is equal to the cross ratio of the parameters defining the involutions to which they belong.*

9. **Corresponding elements under the class.** Let X be a point of the null space $\mathfrak{M}$ of M. Then, $AX = BX = Y$ (say). Hence also $CX = Y$. If $Y = \pm X$, X is a fixed or self-corresponding point under every transformation of the class. If $Y \neq \pm X$, then $AY = BY = CY = X$. Such a pair of points X, Y is called a *pair of corresponding points under the class.* Similarly, a plane U of the null space $\mathfrak{M}'$ is either a self-corresponding plane or one of a pair of corresponding planes under the class.

A pair of lines m, n such that points on m correspond to points on n under every transformation of the class is called a pair of corresponding lines under the class. Evidently, a line containing two points of $\mathfrak{M}$ is self-corresponding.

Let X be a point not in $\mathfrak{M}$. If then $AX = Y$, $BX = Z \neq Y$; and $AY = BZ = X$. Let $AZ = Z'$ and $BY = Y'$; i. e., $ABX = Z'$, $BAX = Y'$. On adding these last two equations, member for member, and recalling (3) we have

$$Y' + Z' = 2X.$$

Hence, Y' and Z' lie on a line through X. Moreover, under any transformation of the class a point on this line m corresponds to a point on the line n through Y and Z. Since X is not a point of $\mathfrak{M}$ there is on m a unique point, in fact the point $X' = X - Y'$, of $\mathfrak{M}$, and moreover, Y' and Z' separate X and X' harmonically.

We therefore have the following theorem.

THEOREM VII. *If A and B determine a linear class and if X is any point in space not in $\mathfrak{M}$, there is through X a unique line m which is one of a pair of corresponding lines m, n under the class. On m there is a single point X' of $\mathfrak{M}$. If then $ABX = Z'$, $BAX = Y'$, Y' and Z' lie on m and separate X and X' harmonically.*

10. **Sufficient conditions that A and B may determine a linear class.** Let both A and B have characteristic roots of multiplicity p at $+1$. Suppose further that the points of the $p-1$ spaces S_A, S_B arise from the same characteristic root. Otherwise we may merely change the sign of one of the matrices.

If S_A coincides with S_B by Corollary 2, Theorem IV, $r' = 0$; i. e., $E = (A-I)(B+I) = -(B-I)(A+I) = 0$. On multiplying out and subtracting, we have, $AB + BA = 2I$, so that A and B determine a linear class. Similarly, if T_A and T_B coincide A and B determine a linear class.

Otherwise, suppose that A and B have in common

(1) $\sigma+\tau$ $(\sigma+\tau \geqq p;\ \geqq q = n-p)$ fixed points of which Y_i $(i = 1, \cdots, \sigma < p)$ are common points of S_A and S_B, and Z_i $(i = 1, \cdots, \tau < q)$ are common points of T_A and T_B.

(2) $\sigma+\tau$ fixed planes of which U_i $(i = 1, \cdots, p-q+\tau)$ are common planes of S'_A and S'_B, and such that U_i $(i = 1, \cdots, p-\sigma)$ are on all the points Y and on every point of T_A and of T_B, and of which V_i $(i = 1, \cdots, q-p+\sigma)$ are common planes of T'_A and T'_B and such that V_i $(i = 1, \cdots, q+\tau)$ are on all the points X and on every point of S_A and of S_B.

Then A and B determine a linear class.

We shall show that the matrix $\psi = AB+BA-2I$ is zero by showing that there exist n linearly independent matrices X such that $\psi X = 0$. We may choose bases of the spaces S_A, S_B, etc., as follows:

$$S_A:\ (Y_1, \cdots, Y_\sigma, Y'_{\sigma+1}, \cdots, Y'_p); \qquad S_B:\ (Y_1, \cdots, Y_\sigma, Y''_{\sigma+1}, \cdots, Y''_p);$$
$$T_A:\ (Z_1, \cdots, Z_\tau, Z'_{\tau+1}, \cdots, Z'_q); \qquad T_B:\ (Z_1, \cdots, Z_\tau, Z''_{\tau+1}, \cdots, Z''_q).$$

If then $X = Y_i$ $(i = 1, \cdots, \sigma)$, or if $X = Z_i$ $(i = 1, \cdots, \tau)$, evidently $\psi X = 0$.

Consider now a point Z''. By hypothesis the $n-p+\sigma$ linearly independent points Y_i $(i = 1, \cdots, \sigma)$, Z_i $(i = 1, \cdots, \tau)$, Z'_i $(i = \tau+1, \cdots, q)$ lie on the $p-\sigma$ linearly independent planes U_j. Hence, a point Z'' lying also on these planes depends linearly on the Y's, the Z's and the (Z')'s. We may write

$$Z'' = c_i Y_i + d_i Z_i + e_i Z'_i,$$

where the repeated index indicates summation over the total number of Y's, etc., in each set. We have,

$$BZ'' = -Z'';$$
$$AZ'' = c_i Y_i - d_i Z_i - e_i Z'_i = 2c_i Y_i - Z''.$$

Hence,

$$BAZ'' = 2c_i Y_i + Z'',$$

and

$$ABZ'' = -AZ'' = -2c_i Y_i + Z'',$$

so that on adding, we have

$$(AB+BA)Z'' = 2Z''.$$

Hence, if $X = Z_i''$ $(i = \tau+1, \cdots, q)$, $\psi X = 0$.

Similarly, if $X = Y_i''$ $(i = \sigma+1, \cdots, p)$, $\psi X = 0$.

Since we have exhibited n linearly independent points X such that $\psi X = 0$, then $\psi = 0$, so that by (3) A and B determine a linear class.

11. **Involutions in the plane** $(n = 3)$. Here $p = 2$ and the ranks of M, E and F are 1, 0 and 1 (or 1, 1 and 0), respectively. A and B are harmonic homologies with axes a and b and centers P_A and P_B, respectively. It follows at once from the general theory that A and B determine a linear class if, and only if, a coincides with b, or by duality, P_A coincides with P_B. In the former case let a and b coincide in l. Then the axis c of C also coincides with l, while the center P_C of C lies on the line $P_A P_B$. Every point on $P_A P_B$ except Q, its point of intersection with l, is the center of some homology of the class, and the cross ratio of any four of these points is equal to the cross ratio of the parameters defining the four homologies to which they belong.

The only pairs of corresponding points under the class are points on l and these points are self-corresponding. A line m through Q corresponds under every involution of the class to its harmonic conjugate n with respect to l and $P_A P_B$. There are no other pairs of corresponding lines.

Let X be any point in the plane not on l or $P_A P_B$. If AB carries X to Y and BA carries X to Z, then Y and Z lie on the line m joining X to Q and separate X and Q harmonically.

If A and B have a common center, the situation is the dual of that described.

12. **Point-plane reflections** $(n = 4)$. If $p = 3$ A and B are point-plane reflections with planes of fixed points π_A, π_B and points of fixed planes P_A, P_B, respectively. The situation is the space analog of that described in § 11. A and B determine a linear class if, and only if, π_A coincides with π_B or P_A coincides with P_B.

In the former case, i. e., if π_A coincides with π_B in π, all points in π are self-corresponding, and there are no other pairs of corresponding points. If the line $P_A P_B$ meets π in Q, any plane through Q corresponds under the class to its harmonic conjugate with respect to π and that plane τ of the pencil which passes through $P_A P_B$. There are no other pairs of corresponding planes.

$P_A P_B$ and lines in π are self-corresponding lines under the class. Any line m through Q determines with $P_A P_B$ a plane α and m corresponds to

its harmonic conjugate in α with respect to $P_A P_B$ and the line of intersection of α and π. There are no other pairs of corresponding lines.

13. **Linear classes of line reflections.** If $p = 2$, A and B are line reflections with axes a, a' and b, b', respectively. It then follows directly from the general theory that A and B determine a linear class if, and only if, either they have an axis in common or a and b lie in a plane π and meet at P, a' and b' lie in a plane π' and meet at P' and PP' is the line of intersection of π and π'.

14. **Line reflections with an axis in common and whose other axes intersect.** Here $p = 2$ and the ranks of M, E, and F are 1, 0 and 1, (or 1, 1 and 0) respectively. In this case the axes a of A and b of B which arise from the characteristic root $+1$ coincide in l. a' and b' lie in a plane π and intersect at P. As the parameter defining C varies c coincides with a and b in l while c' lies in π and turns about the point P. If τ is the plane of l and P points in τ are transformed under the class as under a harmonic homology with axis l and center P. Thus every point in τ is one of a pair of corresponding points under the class, and there are no other such pairs. If π meets l at Q any plane through Q is one of a pair of corresponding planes under the class, and there are no other such pairs. Any line in π is a self-corresponding line under the class. Any line in τ or through Q is one of a pair of corresponding lines under the class and there are no other such pairs.

15. **Line reflections with an axis in common and whose other axes do not intersect.** Here $p = 2$ and the ranks of M, E and F are 2, 0 and 2 (or 2, 2 and 0) respectively. In this case a and b coincide in l while a' and b' do not intersect. c coincides with a and b in l and as the parameter defining C varies, c' sweeps out a quadric surface of which l, a' and b' are lines of a regulus.

Points (planes) on l are self-corresponding and there are no other pairs of corresponding points (planes).

Any line of the regulus of which l, a' and b' are directrices is a self-corresponding line under the class. The only pairs of corresponding lines under the class are tangents m, n to the surface at a point P of l and m and n are harmonic conjugates with respect to the two rulings through P.

16. **Line reflections without an axis in common.** Here $p = 2$ and the ranks of M, E and F are 2, 1 and 1, respectively. In this case a and b lie in a plane π and intersect at a point P; a' and b' lie in a plane π' and intersect at a point P' (and PP' is the line of intersection of π and π'). As the parameter defining C varies c lies in π and turns about P, while c' lies in π and turns about P', and the cross ratio of any four lines c is equal to the cross ratio of the corresponding lines c', and this in turn is

equal to the cross ratio of the parameters defining the four involutions C to which these lines belong.

The only pairs of corresponding points (planes) under the class are points (planes) on the line PP' and each point (plane) on this line corresponds to its harmonic conjugate with respect to $P(\pi)$ and $P'(\pi')$.

Any line in π (π') and passing through P' (P) is a self-corresponding line under the class. If X is any point in space not on π or π' there is through X a unique line m which is one of a pair of corresponding lines m, n under the class. These lines m, n lie in planes τ, τ' of the pencil PP' which are harmonic conjugates with respect to π and π'. Moreover, there is a definite point S at which the line m through X meets PP' and n meets PP' at a point S' which is the harmonic conjugate of S with respect to P and P'.

THE UNIVERSITY OF NORTH CAROLINA.

EUCLIDEAN INVARIANTS OF PLANE ALGEBRAIC CURVES.*

By Louis Weisner.†

I. Introduction.

In a series of fifteen articles published in the *Wiener Sitzungsberichte* between 1913 and 1919, R. Weitzenböck laid the foundations of the invariant theory of forms with respect to the euclidean group and applied his results to the determination of complete systems of particular cases. While cartesian coördinates, which Weitzenböck employs, are appropriate in an analytic treatment of euclidean geometry, minimal coördinates are better adapted to the study of *plane* euclidean geometry and will be used throughout this paper.

The euclidean group assumes the analytic form

$$\begin{aligned} &(1a) \qquad x_1 = \lambda_1 x_1' + \mu_1 x_3', \quad x_2 = \lambda_2 x_2' + \mu_2 x_3', \quad x_3 = x_3' \\ &(1b) \qquad x_1 = \lambda_1 x_2' + \mu_1 x_3', \quad x_2 = \lambda_2 x_1' + \mu_2 x_3', \quad x_3 = x_3' \end{aligned} \qquad (\lambda_1 \lambda_2 \neq 0)$$

in homogeneous minimal point coördinates. Denoting by ξ_1, ξ_2, ξ_3 the homogeneous minimal coördinates of a line in the plane, we find the equations (1) induce the transformations

$$(2a) \qquad \xi_1 = \frac{\xi_1'}{\lambda_1}, \quad \xi_2 = \frac{\xi_2'}{\lambda_2}, \quad \xi_3 = -\frac{\mu_1}{\lambda_1}\xi_1' - \frac{\mu_2}{\lambda_2}\xi_2' + \xi_3',$$

$$(2b) \qquad \xi_1 = \frac{\xi_2'}{\lambda_1}, \quad \xi_2 = \frac{\xi_1'}{\lambda_2}, \quad \xi_3 = -\frac{\mu_2}{\lambda_1}\xi_1' - \frac{\mu_1}{\lambda_2}\xi_2' + \xi_3'.$$

The transformations (1) and (2) together define the euclidean group. Equations (1a) and (2a) define the *direct* euclidean group, which is generated by the direct group of translations

$$(3) \qquad \begin{aligned} &x_1 = x_1' + \mu_1 x_3', \quad x_2 = x_2' + \mu_2 x_3', \quad x_3 = x_3', \\ &\xi_1 = \xi_1', \quad \xi_2 = \xi_2', \quad \xi_3 = -\mu_1 \xi_1' - \mu_2 \xi_2' + \xi_3', \end{aligned}$$

and the direct group of rotations and similarity transformations

$$(4) \qquad \begin{aligned} &x_1 = \lambda_1 x_1', \quad x_2 = \lambda_2 x_2, \quad x_3 = x_3' \\ &\xi_1 = \frac{\xi_1'}{\lambda_1}, \quad \xi_2 = \frac{\xi_2'}{\lambda_2}, \quad \xi_3 = \xi_3' \end{aligned} \qquad (\lambda_1 \lambda_2 \neq 0).$$

* Received December 6, 1927.

† National Research Fellow.

The *entire* euclidean group, group of translations, and group of rotations and similarity transformations are generated by the respective direct groups and the reflection

$$(5) \qquad \begin{aligned} x_1 &= x_2', \quad & x_2 &= x_1', \quad & x_3 &= x_3', \\ \xi_1 &= \xi_2', \quad & \xi_2 &= \xi_2', \quad & \xi_3 &= \xi_3'. \end{aligned}$$

When cartesian coördinates are employed it is necessary to consider more or less independently invariants with respect to the group of rotations and invariants with respect to the group of translations. One of the advantages gained by the introduction of minimal coördinates is due to the simple analytic form which the group of rotations and similarity transformations then assumes, in virtue of which it is possible to determine by inspection whether or not a given expression is an invariant of one or more ternary forms with respect to this group. The theory of euclidean invariants is reduced to the theory of invariants with respect to the group of translations.

We shall confine our attention to the direct euclidean group. We shall determine two fundamental systems of concomitants of a set of ternary forms, in terms of each of which all others can be expressed rationally but not integrally. One of the fundamental systems leads to a set of *absolute* invariants which furnish a criterion for the equivalence of two algebraic curves with respect to the euclidean group and are useful in classifying curves of a given order or class. The methods are similar to those employed in the theory of seminvariants of binary forms.

II. Invariants with respect to rotations and similarity transformations.

1. **Notation.** We write the general algebraic curve of order n in the form

$$(6) \qquad (ax)^n \equiv \sum_{p,q,r} \frac{n!}{p!\,q!\,r!}\, a_{pq}\, x_1^p\, x_2^q\, x_3^r = 0, \quad (p+q+r=n),$$

and the general algebraic curve of class m in the form

$$(7) \qquad (\alpha\xi)^m \equiv \sum_{p,q,r} \frac{m!}{p!\,q!\,r!}\, \alpha_{pq}\, \xi_1^p\, \xi_2^q\, \xi_3^r = 0, \quad (p+q+r=m),$$

where

$$(ax) \equiv a_1 x_1 + a_2 x_2 + a_3 x_3, \quad (\alpha\xi) = \alpha_1 \xi_1 + \alpha_2 \xi_2 + \alpha_3 \xi_3,$$
$$a_1^p\, a_2^q\, a_3^r = a_{pq} \quad (p+q+r=n), \quad \alpha_1^p\, \alpha_2^q\, \alpha_3^r = \alpha_{pq} \quad (p+q+r=m).$$

We shall generally be concerned with a number of groundforms $(ax)^{n_1}$, $(bx)^{n_2}, \cdots, (\alpha\xi)^{m_1}, (\beta\xi)^{m_2}, \cdots$, the coefficients of a point curve being always

denoted by Roman letters and the coefficients of a line curve by Greek letters.

2. **Indices.** The transformation (3) induces the transformations

$$a'_{pq} = \lambda_1^p \lambda_2^q a_{pq}, \qquad \alpha'_{pq} = \lambda_1^{-p} \lambda_2^{-q} \alpha_{pq}$$

on the coefficients of $(ax)^n$, $(\alpha\xi)^m$ respectively. Hence

$$T \equiv x_1^{t_1} x_2^{t_2} x_3^{t_3} \xi_1^{\tau_1} \xi_2^{\tau_2} \xi_3^{\tau_3} \prod_{p,q} a_{pq}^{r_{pq}} b_{pq}^{s_{pq}} \cdots \alpha_{pq}^{\varrho_{pq}} \beta_{pq}^{\sigma_{pq}}$$

is transformed by (3) into

$$T' = \lambda_1^{i_1} \lambda_2^{i_2} T,$$

where

$$i_1 = \tau_1 - t_1 + \sum_{p,q} p(r_{pq} + s_{pq} + \cdots - \varrho_{pq} - \sigma_{pq} - \cdots),$$

$$i_2 = \tau_2 - t_2 + \sum_{p,q} q(r_{pq} + s_{pq} + \cdots - \varrho_{pq} - \sigma_{pq} - \cdots).$$

These two ordered numbers are the *indices* of T. A concomitant with respect to the group of rotations and similarity transformations must be *equi-indical*; that is, every term must have the same indices. The indices of a monomial equal the sum of the indices of its factors; $a_{pq}, b_{pq}, \cdots$ having indices p, q; $\alpha_{pq}, \beta_{pq}, \cdots$ indices $-p, -q$; x_1 indices $-1, 0$; ξ_1 indices $1, 0$; x_2 indices $0, -1$; ξ_2 indices $0, 1$; x_3 and ξ_3 indices $0, 0$. It being possible to determine by inspection whether or not a polynomial is equi-indical, it only remains to consider whether it is an invariant with respect to the direct group of translations. If it is, it is an invariant with respect to the direct euclidean group.

III. First fundamental system.

3. **Annihilators.** Four of the six annihilators of projective invariants of a ternary form are annihilators of euclidean invariants. They are:

$$\Omega_1(\alpha) \equiv \sum_{p,q} p\, \alpha_{p-1,q} \frac{\partial}{\partial \alpha_{pq}},$$

$$\Omega_2(\alpha) \equiv \sum_{p,q} q\, \alpha_{p,q-1} \frac{\partial}{\partial \alpha_{pq}},$$

$$\Omega_3(a) \equiv \sum_{p,q} (n-p-q)\, a_{p+1,q} \frac{\partial}{\partial a_{pq}},$$

$$\Omega_4(a) \equiv \sum_{p,q} (n-p-q)\, a_{p,q+1} \frac{\partial}{\partial a_{pq}}.$$

As indicated by the notation, the first two operators annihilate invariants of a line curve while the last two annihilate invariants of a point curve.

The translation

$$x_1 = x_1' + \mu x_3', \quad x_2 = x_2', \quad x_3 = x_3'$$
$$\xi_1 = \xi_1', \quad \xi_2 = \xi_2', \quad \xi_3 = \xi_3' - \mu \xi_1',$$

where μ is an infinitesimal, induces the infinitesimal transformations

$$a'_{pq} = a_{pq} + (n-p-q)\,\mu\, a_{p+1,\,q}$$
$$\alpha'_{pq} = \alpha_{pq} - p\,\mu\, \alpha_{p-1,\,q}$$

on the coefficients of $(ax)^n$, $(\alpha\xi)^m$ respectively. It follows* that every concomitant of the ground forms is annihilated by

$$\xi_1 \frac{\partial}{\partial \xi_3} - x_3 \frac{\partial}{\partial x_1} + \Omega_3(a) + \Omega_3(b) + \cdots - \Omega_1(\alpha) - \Omega_1(\beta) - \cdots,$$

and similarly by

$$\xi_2 \frac{\partial}{\partial \xi_3} - x_3 \frac{\partial}{\partial x_2} + \Omega_4(a) + \Omega_4(b) + \cdots - \Omega_2(\alpha) - \Omega_2(\beta) - \cdots.$$

These annihilators will be briefly denoted by

$$\xi_1 \frac{\partial}{\partial \xi_3} + \Omega$$

and

$$\xi_2 \frac{\partial}{\partial \xi_3} + \bar{\Omega}$$

respectively.

4. **Concomitant determined by its source.** The conditions that

$$K \equiv \sum_{p,q} \frac{\omega!}{b!\, q!\, (\omega-p-q)!}\, \varphi_{pq}\, x_1^p\, x_2^q\, x_3^{\omega-p-q}$$

be a concomitant of the ground forms are

$$\left(\xi_1 \frac{\partial}{\partial \xi_3} + \Omega\right) \varphi_{pq} = (\omega-p-q)\ \varphi_{p+1,\,q}$$
$$\left(\xi_2 \frac{\partial}{\partial \xi_3} + \bar{\Omega}\right) \varphi_{pq} = (\omega-p-q)\ \varphi_{p,\,q+1}.$$

The concomitant is therefore completely determined by its *source*, φ_{00} up to a factor which is a power of x_3, and may be written

$$K = x_3^q \sum_{p,q} \left(\xi_1 \frac{\partial}{\partial \xi_3} + \Omega\right)^p \left(\xi_2 \frac{\partial}{\partial \xi_3} + \bar{\Omega}\right)^q \frac{\varphi_{00}}{p!\, q!}\, x_1^p\, x_2^q\, x_3^{\omega-p-q}.$$

* Cf. Dickson, Algebraic Invariants, p. 42.

5. **Fundamental concomitants.** In particular, each coefficient of the ground forms is the source of some covariant. We find that a_{rs} and α_{rs} are the sources of

$$A_{rs} \equiv \sum_{p,q} \frac{(n-r-s)!}{(n-r-s-p-q)!\,p!\,q!}\, a_{r+p,s+q}\, x_1^p\, x_2^q\, x_3^{n-p-q-r-s}$$

$$A'_{rs} \equiv \sum_{p,q} (-1)^{p+q} \binom{r}{p} \binom{s}{q} \alpha_{r-p,s-q}\, x_1^p\, x_2^q\, x_3^{m-p-q}$$

respectively. Similarly b_{rs}, β_{rs}, $\cdots$ are the sources of concomitants, B_{rs}, B'_{rs}, $\cdots$. A_{rs} is obtained by polarising $(ax)^n$ r times with respect to the circular point (100) and s times with respect to the circular point (010). In the symbolic notation

$$A_{rs} = a_1^r\, a_2^s\, (ax)^{n-r-s}$$

$$A'_{rs} = \alpha_3^{n-r-s}\, (\alpha_3 x_1 - \alpha_1 x_3)^r\, (\alpha_3 x_2 - \alpha_2 x_3)^s.$$

The concomitants of which ξ_1, ξ_2 are the sources are ξ_1, ξ_2 themselves, while ξ_3 is the source of

$$(\xi x) \equiv \xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3.$$

THEOREM. *The concomitants of which the equi-indical polynomial* $\varphi(\xi_1, \xi_2, \xi_3, a_{pq}, b_{pq}, \cdots, \alpha_{pq}, \beta_{pq}, \cdots)$ *is the source are*

$$K \equiv x_3^g\, \varphi[\xi_1, \xi_2, (\xi x), A_{pq}, B_{pq}, \cdots, A'_{pq}, B'_{pq}, \cdots]$$

where g *is* 0 *or an arbitrary integer greater than some fixed negative integer.*

K is a concomitant, being equi-indical and annihilated by $\xi_1 \dfrac{\partial}{\partial \xi_3} + \Omega$ and $\xi_2 \dfrac{\partial}{\partial \xi_3} + \bar{\Omega}$. Its source, which is the coefficient of the highest power of x_3, is clearly $\varphi(\xi_1, \xi_2, a_{pq}, b_{pq}, \cdots, \alpha_{pq}, \beta_{pq}, \cdots)$. It may happen that the expanded form of $\varphi[\xi_1, \xi_2, (\xi x), A_{pq}, B_{pq}, \cdots, A'_{pq}, B'_{pq}, \cdots]$ is divisible by x_3^h $(h > 0)$. If so, K is a rational integral concomitant for every integral value of $g \geqq -h$.

Cor. 1. *Every rational integral concomitant of the ground forms equals a rational integral function of* $x_3, \xi_1, \xi_2, (\xi x), A_{pq}, B_{pq}, \cdots, A'_{pq}, B'_{pq}, \cdots$, *divided by a power of* x_3.

Cor. 2. *A syzygy* $f(K_1, \cdots, K_r) = 0$ *among the concomitants of a set of forms implies that* $f(S_1, \cdots, S_r) = 0$, S_i *being the source of* K_i.

6. **Complete system of a conic.** The results of the preceding section are useful in determining whether a given concomitant equals a rational integral function of certain other concomitants. If $K_1, \cdots, K_r$ are con-

comitants of the ground forms and $K_1 = P(K_2, \cdots, K_r)$, P being a polynomial in its arguments, then $S_1 = P(S_2, \cdots, S_r)$, S_i being the source of K_i. It is a relatively simple matter to determine whether S_1 equals a rational integral function of $S_2, \cdots, S_r$. If we find that $S_1 = P(S_2, \cdots, S_r)$, we infer that $K_1 = x_3^g P(K_2, \cdots, K_r)$, g being determined by the fact that the equation in which it appears is homogeneous in x_1, x_2, x_3. Hence $K_1 = P(K_2, \cdots, K_r)$, if homogeneous in x_1, x_2, x_3, is a consequence of $S_1 = P(S_2, \cdots, S_r)$.

On applying this method to the complete system of the point conic and of the line conic derived by Weitzenböck,* I find that his 15 concomitants of the point conic form an irreducible system, while his 24 concomitants of the line conic contain two contravariants and one mixed form which are reducible. The reducibility of these concomitants is demonstrated by the following equations which are written in Weitzenböck's notation:

$$(\alpha\beta p)(\alpha u')(\beta l') = -\frac{i}{4}(\alpha\beta q)^2(u'q) + \frac{i}{4}(\alpha\beta p)(\alpha\beta q)(u'p),$$

$$(\alpha\beta q)(\alpha u')(\beta l') = \frac{i}{4}(\alpha\beta p)^2(u'q) - \frac{i}{4}(\alpha\beta p)(\alpha\beta q)(u'q),$$

$$(\alpha\beta x)(\alpha u')(\beta l') = \frac{i}{4}(u'q)(\alpha\beta x)(\alpha\beta p) - \frac{i}{4}(u'p)(\alpha\beta x)(\alpha\beta q).$$

IV. Second fundamental system.

7. Single point curve. The derivatives of $(ax)^n$ with respect to x_1 and x_2 are covariants, as $\frac{\partial^{r+s}(ax)^n}{\partial x_1^r \partial x_2^s}$ is the result of polarising $(ax)^n$ r times with respect to the circular point (100) and s times with respect to the circular point (010). In particular we have the covariant lines

$$P_{n-1,0} \equiv \frac{\partial^{n-1}}{\partial x_1^{n-1}}(ax)^n = a_{n0}x_1 + a_{n-1,1}x_2 + a_{n-1,0}x_3 = 0$$

$$P_{0,n-1} \equiv \frac{\partial^{n-1}}{\partial x_2^{n-1}}(ax)^n = a_{1,n-1}x_1 + a_{0n}x_2 + a_{0,n-1}x_3 = 0,$$

which intersect in the contravariant point

$$Q \equiv \Delta_1\xi_1 + \Delta_2\xi_2 + \Delta_3\xi_3 = 0,$$

where

$$\Delta_1 \equiv a_{n-1,1}a_{0,n-1} - a_{n-1,0}a_{0n}$$

$$\Delta_2 \equiv a_{n-1,0}a_{1,n-1} - a_{n0}a_{0,n-1}$$

$$\Delta_3 \equiv a_{n0}a_{0n} - a_{n-1,1}a_{1,n-1}.$$

* R. Weitzenböck, *Über relative Bewegungsinvarianten*, Math. Annalen, vol. 93, 1925, p. 58.

I call this point the *pole* of the curve. It is the unique point whose first polar with respect to the curve passes through both circular points. It exists for all curves of order $n>1$ for which Δ_1, Δ_2, Δ_3 are not simultaneously zero.

On translating the pole of $(ax)^n$ to the origin, we find that, if $\frac{n!}{r!\,s!\,(n-r-s)!}\,a'_{rs}$ denotes the new coefficient of $x_1^r x_2^s x_3^{n-r-s}$, then

$$I_{rs} \equiv \Delta_3^{n-r-s} a'_{rs} = \sum_{p=q,r=s}^{p+q=n} \frac{(n-r-s)!}{(p-r)!\,(q-s)!\,(n-p-q)!}\, a_{pq}\, \Delta_1^{p-r}\, \Delta_2^{q-s}\, \Delta_3^{n-p-q}.$$

In particular, $I_{n-1,0} = I_{0,n-1} = 0$. *The I's are invariants of* $(ax)^n$, being equi-indical and annihilated by $\Omega_3(a)$ and $\Omega_4(a)$.

THEOREM. *Every rational integral concomitant of* $(ax)^n$ *equals a rational integral function of* $P_{n-1,0}$, $P_{0,n-1}$, Q, ξ_1, ξ_2, x_3, *and the I's, divided by a power of* Δ_3.

Applying the particular translation

$$x_1 = x_1' + \frac{\Delta_1}{\Delta_3} x_3', \quad x_2 = x_2' + \frac{\Delta_2}{\Delta_3} x_3', \quad x_3 = x_3'$$

$$\xi_1 = \xi_1', \quad \xi_2 = \xi_2', \quad \xi_3 = -\frac{\Delta_1}{\Delta_3}\xi_1' - \frac{\Delta_2}{\Delta_3}\xi_2' + \xi_3'$$

(which carries the pole of $(ax)^n$ to the origin) to the concomitant $K \equiv K(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3, a_{pq})$, we must have, by definition of *concomitant*,

$$K = K\left(x_1 - \frac{\Delta_1}{\Delta_3} x_3,\ x_2 - \frac{\Delta_2}{\Delta_3} x_3,\ x_3,\ \xi_1,\ \xi_2,\ \frac{Q}{\Delta_3},\ \frac{I_{pq}}{\Delta_3^{n-p-q}}\right).$$

The theorem follows from the facts that

$$\Delta_3 x_1 - \Delta_1 x_3 = a_{0n} P_{n-1,0} - a_{n-1,1} P_{0,n-1}$$
$$\Delta_3 x_2 - \Delta_2 x_3 = a_{n0} P_{0,n-1} - a_{1,n-1} P_{n-1,0}$$
$$I_{pq} \equiv a_{pq} \quad (p+q = n).$$

8. **Single line curve.** The linear polar of l_∞ with respect to $(a\xi)^m = 0$, obtained by differentiating this equation $m-1$ times with respect to ξ_3, is the point

$$G \equiv \alpha_{10}\xi_1 + \alpha_{01}\xi_2 + \alpha_{00}\xi_3 = 0.$$

The name *center* has been proposed for this point and will be adopted here.* The minimal lines through the center are

$$L_1 \equiv \alpha_{00} x_1 - \alpha_{10} x_3 = 0, \quad L_2 \equiv \alpha_{00} x_2 - \alpha_{01} x_3 = 0.$$

* Cf. H. M. Taylor, *On the center of an algebraical curve*, Quart. Jour., vol. 24, p. 55; H. F. Baker, *ibid.*, p. 338.

On translating the center to the origin, we find that, if α'_{rs} denotes the new value of α_{rs}, then

$$J_{rs} \equiv \alpha_{00}^{r+s-1}\,\alpha'_{rs} = \sum_{p,\,q=0}^{p=r,\,q=s} (-1)^{r+s-p-q} \binom{r}{p}\binom{s}{q} \alpha_{00}^{p+q-1}\,\alpha_{10}^{r-p}\,\alpha_{01}^{s-q}\,\alpha_{pq}.$$

In particular, $J_{10} = J_{01} = 0$. *The J's are invariants of* $(\alpha\xi)^m$. As in § 7, we have the

THEOREM. *Every rational integral concomitant of* $(\alpha\xi)^m$ *equals a rational integral function of* $L_1, L_2, G, \xi_1, \xi_2, x_3$ *and the J's, divided by a power of* α_{00}.

9. **Several curves.** Consider next the fundamental system of $(ax)^{n_1}$, $(bx)^{n_2}, \cdots, (\alpha\xi)^{m_1}, (\beta\xi)^{m_2}, \cdots$. If at least one line curve, say $(\alpha\xi)^{m_1}$, is present in the set, translate the entire set of curves so that the center of $(\alpha\xi)^{m_1}$ is at the origin. Then α_{pq} becomes the α'_{pq} of § 8; let $a_{pq}, b_{pq}, \cdots$ $\cdots, \beta_{pq}, \cdots$ become $a''_{pq}, b''_{pq}, \cdots, \beta''_{pq}, \cdots$.

To calculate β''_{pq} we must translate $(\beta\xi)^{m_2}$ so that the center of $(\alpha\xi)^{m_1}$ is at the origin. We do so in two steps. First translate the center of $(\beta\xi)^{m_2}$ to the origin, so that β_{pq} becomes β'_{pq} (which bears the same relation to β_{pq} that α'_{pq} bears to α_{pq}). The translation

$$x_1 = x'_1 + \left(\frac{\beta_{10}}{\beta_{00}} - \frac{\alpha_{10}}{\alpha_{00}}\right) x'_3, \qquad x_2 = x'_2 + \left(\frac{\beta_{01}}{\beta_{00}} - \frac{\alpha_{01}}{\alpha_{00}}\right) x'_3, \qquad x_3 = x'_3$$

then carries the center of $(\alpha\xi)^{m_1}$ to the origin. Hence β''_{pq} equals a rational function of the β's, $\alpha_{00}\beta_{10} - \alpha_{10}\beta_{00}$, $\alpha_{00}\beta_{01} - \alpha_{01}\beta_{00}$ and α_{00}, all of which are invariants.

Similarly, we find that a''_{pq} equals a rational function of the a''s of § 7, $\alpha_{00}\Delta_1 - \alpha_{10}\Delta_3$, $\alpha_{00}\Delta_2 - \alpha_{01}\Delta_3$ and α_{00}. This result is meaningless if $m_1 = 1$, in which case the pole $(\Delta_1, \Delta_2, \Delta_3)$ does not exist. But then we find by direct calculations that

$$a''_{10} = a_{10}, \quad a''_{01} = a_{01}, \quad a''_{00} = \alpha_{10}a_{10} + \alpha_{01}a_{01} + \alpha_{00}a_{00},$$

all of which are invariants.

The concomitant $K = K(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3, a_{pq}, b_{pq}, \cdots, \alpha_{pq}, \beta_{pq}, \cdots)$ of the ground forms is invariant under the translation

$$x_1 = x'_1 + \frac{\alpha_{10}}{\alpha_{00}}, \qquad x_2 = x'_2 + \frac{\alpha_{01}}{\alpha_{00}}, \qquad x_3 = x'_3,$$

$$\xi_1 = \xi'_1, \qquad \xi_2 = \xi'_2, \qquad \xi_3 = -\frac{\alpha_{10}}{\alpha_{00}}\xi'_1 - \frac{\alpha_{01}}{\alpha_{00}}\xi'_2 + \xi'_3$$

which carries the center of $(\alpha\xi)^{m_1}$ to the origin. Hence

$$K = K\left(x_1 - \frac{\alpha_{10}}{\alpha_{00}} x_3,\ x_2 - \frac{\alpha_{01}}{\alpha_{00}} x_3,\ x_3,\ \xi_1,\ \xi_2,\ \frac{G}{\alpha_{00}},\ a''_{pq},\ b''_{pq},\ \cdots,\ \alpha'_{pq},\ \beta''_{pq},\ \cdots\right),$$

where $G = 0$ is the center of $(\alpha\xi)^{m_1} = 0$. *Hence K equals a rational function of the fundamental system of $(\alpha\xi)^{m_1}$, the fundamental invariants of each of the other forms, and certain linear joint invariants of $(\alpha\xi)^{m_1}$ and each of the other forms.*

If the given set of forms represent only point-curves, at least one of which is non-linear, similar results are obtained on translating the entire set of curves so that the pole of some one of them coincides with the origin.

Finally, to find the fundamental system of a set of linear forms (ax), (bx), (cx), $\cdots$, apply the translation

$$x_1 = x'_1 + \frac{a_{01}b_{00} - a_{00}b_{01}}{a_{10}b_{01} - a_{01}b_{10}} x'_3,$$

$$x_2 = x'_2 + \frac{a_{00}b_{10} - a_{10}b_{00}}{a_{10}b_{01} - a_{01}b_{10}} x'_3,$$

$$x_3 = x'_3, \qquad \xi_1 = \xi'_1, \qquad \xi_2 = \xi'_2,$$

$$\xi_3 = -\frac{a_{01}b_{00} - a_{00}b_{01}}{a_{10}b_{01} - a_{01}b_{10}} \xi'_1 - \frac{a_{00}b_{10} - a_{10}b_{00}}{a_{10}b_{01} - a_{01}b_{10}} \xi'_2 + \xi'_3$$

(which carries the point of intersection of $(ax) = 0$ and $(bx) = 0$ to the origin) to each of these forms, obtaining

$$\begin{aligned}
a'_{10}x_1 + a'_{01}x_2 + a'_{00}x_3 &= a_{10}x_1 + a_{01}x_2 \\
b'_{10}x_1 + b'_{01}x_2 + b'_{00}x_3 &= b_{10}x_1 + b_{01}x_2 \\
c'_{10}x_1 + c'_{01}x_2 + c'_{00}x_3 &= c_{10}x_1 + c_{01}x_2 \begin{vmatrix} a_{10} & a_{01} & a_{00} \\ b_{10} & b_{01} & b_{00} \\ c_{10} & c_{01} & c_{00} \end{vmatrix} x_3
\end{aligned}$$

respectively. All the new coefficients are invariants. If $K = K(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3, a_{pq}, b_{pq}, c_{pq}, \ldots)$ is a concomitant of the linear forms, it is invariant under the preceding translation. We infer that K *equals a rational integral function of*

$$x_3,\ \xi_1,\ \xi_2,\ a_{10},\ a_{01},\ b_{10},\ b_{01},\ c_{10},\ c_{01},\ \begin{vmatrix} a_{10} & a_{01} & a_{00} \\ b_{10} & b_{01} & b_{00} \\ c_{10} & c_{01} & c_{00} \end{vmatrix},$$

$$\begin{vmatrix} -x_3 & 0 & x_1 \\ a_{10} & a_{01} & a_{00} \\ b_{10} & b_{01} & b_{00} \end{vmatrix}, \quad \begin{vmatrix} x_3 & 0 & x_2 \\ a_{10} & a_{01} & a_{00} \\ b_{10} & b_{01} & b_{00} \end{vmatrix}, \quad \begin{vmatrix} \xi_1 & \xi_2 & \xi_3 \\ a_{10} & a_{01} & a_{00} \\ b_{10} & b_{01} & b_{00} \end{vmatrix},$$

divided by a power of $a_{10}\,b_{01} - a_{01}\,b_{10}$. The interpretation of this fundamental system is quite simple.

V. Absolute invariants.

10. **Equivalent point curves.** Two curves are *equivalent* if there exists a transformation of the direct euclidean group which transforms one into the other. Two lines $a_{10}\,x_1 + a_{01}\,x_2 + a_{00}\,x_3 = 0$, $b_{10}\,x_1 + b_{01}\,x_2 + b_{00}\,x_3 = 0$, are equivalent if (1) $a_{10}\,a_{01}\,b_{10}\,b_{01} \neq 0$, in which case neither is a minimal line; (2) $a_{10} = b_{10} = 0$ or $a_{01} = b_{01} = 0$, in which case both are minimal lines passing through the same circular point; (3) $a_{10} = a_{01} = b_{10} = b_{01} = 0$, in which case both equations represent l_∞.

Suppose now that $(a\,x)^n = 0$, $(b\,x)^n = 0$, $(n > 1)$, are equivalent curves whose poles are in the finite part of the plane. Their poles can be made to coincide at the origin by suitable translations, the new coefficients being a'_{pq}, b'_{pq} (§ 7). The curves being still equivalent, there must be a direct euclidean transformation which transforms $(a'x)^n$ into $(b'x)^n$. This transformation must leave the origin fixed since, when two curves coincide, their poles must coincide. The transformation is therefore a rotation and similarity transformation

$$x_1 = \lambda_1\,x'_1, \qquad x_2 = \lambda_2\,x'_2, \qquad x_3 = x'_3 \qquad (\lambda_1\,\lambda_2 \neq 0).$$

In order that this transformation transform

$$(a'\,x)^n \equiv \sum_{p,q,r} \frac{n!}{p!\;q!\;r!}\,a'_{pq}\,x_1^p\,x_2^q\,x_3^r \qquad (p+q+r=n)$$

into

$$(b'\,x)^n \equiv \sum_{p,q,r} \frac{n!}{p!\;q!\;r!}\,b'_{qp}\,x_1^p\,x_2^q\,x_3^r, \qquad (p+q+r=n)$$

it is necessary and sufficient that

$$\lambda_1^p\,\lambda_2^q\,a'_{pq} = k\,b'_{pq},$$

$k \neq 0$ being independent of p, q. In particular

$$a'_{00} = k\,b'_{00}, \qquad \lambda_1\,a'_{10} = k\,b'_{10}, \qquad \lambda_2\,a'_{01} = k\,b'_{01}.$$

Hence

$$\frac{a'^{p+q-1}_{00}\;a'_{pq}}{a'^{p}_{10}\;a'^{q}_{01}} = \frac{b'^{p+q-1}_{00}\;b'_{pq}}{b'^{p}_{10}\;b'^{q}_{01}}.$$

The *absolute* invariants

$$I'_{pq} \equiv \frac{I_{00}^{p+q-1}\,I_{pq}}{I_{10}^{p}\,I_{01}^{q}}$$

must therefore have the same numerical values for the two curves. (p, q) assumes all values consistent with $p+q \leqq n$, excepting $(0, 0)$, $(1, 0)$, $(0, 1)$, $(n-1, 0)$, $(0, n-1)$, in which cases I'_{pq} has the value 1 or 0 for an arbitrary curve of order n. It is necessary to assume $n>2$, since, when $n=2$, $I_{10}=I_{01}=0$. Thus, for every $n>2$, we have a set of $\frac{n(n+3)}{2}-4$ *fundamental* absolute invariants of $(ax)^n$, which, by a counting of constants, is seen to be the minimum number of algebraically independent absolute invariants.

By considering the preceding equations in the reverse order we arrive at the following

THEOREM. *Two curves of order $n>2$ whose poles are in the finite part of the plane are equivalent with respect to the direct euclidean group if their corresponding fundamental absolute invariants have the same numerical values, none of these values being $0, \infty$, or $0/0$.*

In Section VI this theorem will be extended to include conics. By means of it, it is possible to classify all curves of a given order, putting equivalent curves into the same class. The *ordinary* curves of a given order are those whose centers are in the finite part of the plane and whose fundamental absolute invariants are numerically different from $0, \infty, 0/0$. Other curves are *special*, including, among others, degenerate curves and curves which pass through one or both of the circular points. The suggested classification will be carried out in detail for conics.

11. **Equivalent line curves.** By translating the centers of the two curves $(\alpha\xi)^m$, $(\beta\xi)^m$ to the origin, the centers being assumed to be in the finite part of the plane, and considering, as in § 10, the conditions under which the two curves may then be equivalent, we arrive at the absolute invariants

$$J'_{pq} \equiv \frac{J_{00}^{\frac{p+q-2}{2}} J_{pq}}{J_{20}^{p/2} J_{02}^{q/2}},$$

where $(p, q) \neq (0, 0), (1, 0), (0, 1), (2, 0), (0, 2)$. Either value of the two-valued functions $\sqrt{J_{20}}$, $\sqrt{J_{02}}$, $\sqrt{J_{00}}$, may be chosen, but the same values must be taken in every J_{pq}.

A theorem similar to that of § 10 is valid.

VI. The conic.

12. **Absolute invariant of a point conic.** A central conic is one whose pole (which is the same as its center) is in the finite part of the plane. On translating the pole of the conic $(ax)^2$ to the origin, the equation becomes

$$a_{20}\,x_1^2 + 2\,a_{11}\,x_1\,x_2 + a_{02}\,x_2^2 + a'_{00}\,x_3^2 = 0,$$

where

$$a'_{00} = \frac{a_{00}\,a_{20}\,a_{02} - a_{00}\,a_{11}^2 - a_{20}\,a_{01}^2 - a_{10}^2\,a_{02} + 2\,a_{10}\,a_{01}\,a_{11}}{a_{00}\,a_{20} - a_{11}^2},$$

the numerator of this fraction being the discriminant D of the conic. All the coefficients are invariants. The conic has the absolute invariant

$$I(a) = \frac{a_{20}\,a_{02}}{a_{11}^2}.$$

It is readily proved that *the proper central conics* $(ax)^2$, $(bx)^2$, *are equivalent if* $I(a) = I(b) \neq 0,\ \infty,\ 0/0$.

13. **Classification of central point conics.** If $D \neq 0$ and $I(a) \neq 0, \infty, 0/0$, it is possible to determine by the preceding theorem whether $(ax)^2$ is equivalent to some other conic. If one or more of the conditions $D = 0$, $I(a) = 0,\ \infty,\ 0/0$ are satisfied, $(ax)^2$ is a *special* conic. An exhaustive classification of the special conics is easily made, with the following results:

1. $D \neq 0$, $a_{20}\,a_{02} = 0$, not both $a_{20}, a_{02} = 0$. Conic passes through one of the circular points.

2. $D \neq 0$, $a_{20} = a_{02} = 0$. Circle.

3. $D \neq 0$, $a_{11} = 0$. Equilateral hyperbola.

4. $D = 0$, $a_{20}\,a_{02}\,a_{11} \neq 0$. Two ordinary* lines intersecting in the finite part of the plane. If θ denotes one of the angles between the lines,

$$\sec^2\theta = I \equiv \frac{a_{20}\,a_{02}}{a_{11}^2}.$$

Two pairs of ordinary lines which intersect in the finite part of the plane are equivalent if their I's have the same numerical value, this value being different from $0,\ \infty,\ 0/0$.

5. $D = 0$, $a_{20}\,a_{02} = 0$, not both $a_{20}, a_{02} = 0$. A minimal line and an ordinary line.†

6. $D = 0$, $a_{20} = a_{02} = 0$. Two distinct minimal lines intersecting in an ordinary point.

7. $D = 0$, $a_{11} = 0$. Two perpendicular ordinary lines.

---

* An *ordinary* line is one which does not pass through a circular point. An *ordinary* point is one which is in the finite part of the plane.

† The classification being made with respect to the direct euclidean group, this case includes two classes, as a line passing through one circular point is not equivalent to a line passing through the other circular point. The two classes form one class under the entire euclidean group. These remarks apply to some of the subsequent results.

14. **Classification of point parabolas.** A parabola is a conic whose pole does not exist or whose pole, if it exists, is on l_∞. The conic $(ax)^2$ has no pole if the equations (Cf. § 7)

$$a_{11}a_{01} - a_{10}a_{02} = 0, \quad a_{11}a_{10} - a_{01}a_{20} = 0, \quad a_{20}a_{02} - a_{11}^2 = 0,$$

are satisfied simultaneously. Eliminating a_{11}, we obtain

$$\sqrt{a_{02}}\,(a_{01}\sqrt{a_{20}} - a_{10}\sqrt{a_{02}}) = 0,$$
$$\sqrt{a_{20}}\,(a_{10}\sqrt{a_{02}} - a_{01}\sqrt{a_{20}}) = 0.$$

By considering all cases in which these equations may be satisfied we determine all conics which have no poles:

1. $a_{20} = a_{11} = a_{02} = 0$, $a_{10}a_{01} \neq 0$. Ordinary line and l_∞.
2. $a_{20} = a_{11} = a_{02} = a_{10}a_{01} = 0$, but not both $a_{10}, a_{01} = 0$. Minimal line and l_∞.
3. All coefficients* vanish except a_{00}. l_∞ repeated.
4. $a_{10}\sqrt{a_{02}} - a_{01}\sqrt{a_{20}} = 0$, $a_{11}a_{20}a_{02}(a_{00}a_{20} - a_{10}^2) \neq 0$. Two distinct parallel ordinary lines.
5. $a_{10}\sqrt{a_{02}} - a_{01}\sqrt{a_{20}} = a_{00}a_{20} - a_{10}^2 = 0$, $a_{11}a_{20}a_{02} \neq 0$. Two coincident ordinary lines.
6. $a_{20} = a_{11} = a_{10} = 0$, $a_{02}(a_{00}a_{02} - a_{01}^2) \neq 0$; or $a_{02} = a_{11} = a_{01} = 0$, $a_{20}(a_{00}a_{20} - a_{10}^2) \neq 0$. Two distinct minimal lines passing through the same circular point.
7. $a_{20} = a_{11} = a_{10} = a_{00}a_{02} - a_{01}^2 = 0$, $a_{02} \neq 0$; or
   $a_{02} = a_{11} = a_{01} = a_{00}a_{20} - a_{10}^2 = 0$, $a_{20} \neq 0$. Two coincident minimal lines.

The equation of a parabola may be written in the form

$$(\sqrt{a_{20}}\,x_1 + \sqrt{a_{02}}\,x_2)^2 + 2a_{10}x_1x_3 + 2a_{01}x_2x_3 + a_{00}x_3^2 = 0.$$

It is readily proved that if

$$a_{20}a_{02}(a_{10}\sqrt{a_{02}} - a_{01}\sqrt{a_{20}}) \neq 0,$$

we can find a direct euclidean transformation which transforms the equation into

$$(x_1 - x_2)^2 + (x_1 + x_2)x_3 = 0,$$

which is a parabola whose center is (110). The condition

* Throughout the classification it is assumed that not all the coefficients vanish.

$$a_{20}\, a_{02}\, (a_{10} \sqrt{a_{02}} - a_{01} \sqrt{a_{20}}) = 0$$

gives only one new class of special parabolas:

8. $a_{20}\, a_{02} = 0$, but not both $a_{20}, a_{02} = 0$. Parabola through a circular point. The circular point is the pole of the parabola.

15. **Classification of central line conics.** The invariants of the conic $(\alpha\, \xi)^2$ are (Cf. § 8)

$$J_{20} \equiv \alpha_{00}\, \alpha_{20} - \alpha_{10}^2$$
$$J_{11} \equiv \alpha_{00}\, \alpha_{11} - \alpha_{10}\, \alpha_{01}$$
$$J_{02} \equiv \alpha_{00}\, \alpha_{02} - \alpha_{01}^2$$
$$J_{00} \equiv \alpha_{00}.$$

Its discriminant is

$$D \equiv \frac{J_{20}\, J_{02} - J_{11}^2}{J_{00}},$$

and its absolute invariant is

$$J' \equiv \frac{J_{20}\, J_{02}}{J_{11}^2}.$$

Two central conics are equivalent if their J''s have the same value, this value being different from $0, \infty, 0/0$. The case in which $D = 0, J_{00} \neq 0$, presents no exception to this statement, since $J' = 1$; the conic consists of two ordinary points, and any two pairs of ordinary points are equivalent.

The cases in which $J' = 0, \infty, 0/0$ are:

1. $J_{20}\, J_{02} = 0$, $J_{11} \neq 0$, not both $J_{20}, J_{02} = 0$. Conic passes through a circular point.

2. $J_{20} = J_{02} = 0$, $J_{11} \neq 0$. Circle.

3. $J_{11} = 0$, $J_{20}\, J_{02} \neq 0$. Equilateral hyperbola.

4. $J_{11} = 0$, $J_{20}\, J_{02} = 0$, not both $J_{20}, J_{02} = 0$. Two distinct points on a minimal line.

5. $J_{20} = J_{02} = J_{11} = 0$. Two coincident ordinary points.

16. **Classification of line parabolas.** The conic $(\alpha\, \xi)^2$ has no center if $\alpha_{10} = \alpha_{01} = \alpha_{00} = 0$. Its equation then is

$$\alpha_{20}\, \xi_1^2 + 2\, \alpha_{11}\, \xi_1\, \xi_2 + \alpha_{02}\, \xi_2^2 = 0.$$

This conic evidently consists of two points on l_∞. The following cases may occur:

1. $\alpha_{20}\, \alpha_{02}\, \alpha_{11}\, (\alpha_{20}\, \alpha_{02} - \alpha_{11}^2) \neq 0$. Two distinct points on l_∞, neither a circular point.

2. $\alpha_{20}\,\alpha_{02} - \alpha_{11}^2 = 0$. Two coincident points on l_∞, other than a circular point.

3. $\alpha_{20}\,\alpha_{02} = 0$, $\alpha_{11} \neq 0$, not both α_{20}, $\alpha_{02} = 0$. A circular point and another point on l_∞ different from the other circular point.

4. $\alpha_{20}\,\alpha_{02} = \alpha_{11} = 0$. A circular point repeated.

5. $\alpha_{20} = \alpha_{02} = 0$. The circular points.

Every parabola $(\alpha\,\xi)^2$, $\alpha_{00} = 0$, is equivalent to

$$\xi_1\,\xi_2 + \xi_1\,\xi_3 + \xi_2\,\xi_3 = 0$$

if

$$\alpha_{10}\,\alpha_{01}\,D \neq 0,$$

where

$$D \equiv 2\,\alpha_{10}\,\alpha_{01}\,\alpha_{11} - \alpha_{10}^2\,\alpha_{02} - \alpha_{01}^2\,\alpha_{20}$$

is the discriminant of the parabola. The following new special cases occur:

6. $\alpha_{10}\,\alpha_{01} = 0$, $D \neq 0$. Parabola passes through a circular point.

7. $\alpha_{10}\,\alpha_{01} \neq 0$, $D = 0$. An ordinary point and a point on l_∞ other than a circular point.

8. $\alpha_{10}\,\alpha_{01} = 0$, $D = 0$, not both α_{10}, $\alpha_{01} = 0$. An ordinary point and a circular point.

HARVARD UNIVERSITY.

NON-ABELIAN GROUPS WHOSE GROUPS OF ISOMORPHISMS ARE ABELIAN.*

By C. Hopkins.

Introduction. In the appendix of Hilton's *Finite Groups* (1908), page 233, the question whether a non-abelian group can have an abelian group of isomorphisms occurs among "a few interesting questions still awaiting solution". A non-abelian group of order 64 whose group of isomorphisms is abelian and of order 128 was later constructed by G. A. Miller.† No other discussian of the problem appears in the literature of Mathematics. In fact, an exhaustive analysis of the properties of a non-abelian group whose group of isomorphisms is abelian seems to involve some difficulty, for the reason that little is known concerning isomorphisms of a non-abelian group.

The nature of an abelian group whose group of isomorphisms is abelian was determined by G. A. Miller,‡ who proved that *a necessary and sufficient condition that an operation of the group of isomorphisms of an abelian group be invariant under this group is that it should transform every operation of this abelian group into the same power of itself.* From this it is obvious that the only abelian groups whose groups of isomorphisms are abelian are the cyclic groups. In what follows we shall develop a few necessary conditions which must be satisfied by a non-abelian group whose group of isomorphisms is abelian.

1. **The fundamental theorem.** We assume a non-abelian group G which is restricted solely by the hypothesis that its group of isomorphisms I is abelian. Since the central quotient group of G is simply isomorphic with the group of inner isomorphisms of G, it is clear that every commutator of G is invariant under G. The group G is accordingly the direct product of its Sylow subgroups.§ Since every Sylow subgroup of G must correspond to itself in any isomorphism of G with itself, we may confine ourselves to the case where G is of order p^m.

For convenience we shall introduce the following notation: the symbol G shall consistently denote a non-abelian group whose group of isomorphisms I

* Received January 22, 1928.

† G. A. Miller: Mess. of Math., vol. 43 (1913–1914), p. 124.

‡ G. A. Miller: Trans. Amer. Math. Soc., vol. 1 (1900), p. 397; vol. 2 (1901), p. 260.

§ Burnside: Theory of Groups of Finite Order, 1897, p. 115.

is abelian; operations in G and operations of I shall be denoted by lower-case and capital letters respectively; an isomorphism of G with itself will often be called an *automorphism*, for the purpose of distinguishing it from an isomorphism of G with a subgroup of G. The operations $s^{-1}\,t^{-1}\,s\,t$ shall be called a commutator of G when both of its elements are in G; the operations $s^{-1}\,T^{-1}\,s\,T$ shall be referred to as a multiplier of G when one of its elements T is in I. It will be noted that these forms are not mutually exclusive.

Let T_1 and T_2 be two distinct operations of I and let s be any operation of G. Suppose that

$$T_1^{-1}\,s\,T_1 = s\gamma_1 \quad \text{and} \quad T_2^{-1}\,s\,T_2 = s\gamma_2.$$

Also

$$T_1^{-1}\,\gamma_2\,T_1 = \gamma_2\,\gamma_{12} \quad \text{and} \quad T_2^{-1}\,\gamma_1\,T_2 = \gamma_1\,\gamma_{21}.$$

Now

$$T_1^{-1}\,T_2^{-1}\,s\,T_2\,T_1 = T_1^{-1}\,s\gamma_2\,T_1 = s\gamma_1\,\gamma_2\,\gamma_{12},$$

and

$$T_2^{-1}\,T_1^{-1}\,s\,T_1\,T_2 = T_2^{-1}\,s\gamma_1\,T_2 = s\gamma_2\,\gamma_1\,\gamma_{21}.$$

Since I is abelian,

$$T_1\,T_2 = T_2\,T_1.$$

Therefore

$$\gamma_1\,\gamma_2\,\gamma_{12} = \gamma_2\,\gamma_1\,\gamma_{21}.$$

If γ_1 is the identity, then the relation above reduces to

$$\gamma_2\,\gamma_{12} = \gamma_2,$$

which implies that γ_{12} is the identity. From this we have the theorem that *every operation of I which is commutative with an operation s of G is commutative with every conjugate of s under I.*

2. **Commutators and multipliers.** We shall now prove that every commutator of G is a characteristic operation: i. e. an invariant operation under the holomorph of G. Suppose that s and t are any two non-commutative operations of G and let $s^{-1}\,ts = tc$. We shall assume that c is not characteristic: i. e., we shall assume an outer isomorphism T which effects the following correspondences:

$$s \sim ss_0, \qquad t \sim tt_0, \qquad c \sim cc_0.$$

Since conjugates must correspond,

$$s^{-1}\,ts \sim s_0^{-1}\,s^{-1}\,tt_0\,ss_0.$$

Since products must correspond,

$$tc \sim tt_0 cc_0.$$

We assumed that $s^{-1}ts = tc$; therefore

$$s_0^{-1} s^{-1} tt_0 ss_0 = tt_0 cc_0.$$

Now there is an inner isomorphism S in I which transforms the operations of G according to s; in symbols

$$S^{-1} tS = tc.$$

Now

$$S^{-1} T^{-1} tTS = S^{-1} tt_0 S = tct_0 t_{0S},$$

where $t_{0S} = t_0^{-1} S^{-1} t_0 S$, and

$$T^{-1} S^{-1} tST = T^{-1} tcT = tt_0 cc_0.$$

Since I is an abelian,

$$ST = TS.$$

Therefore

$$t_0 cc_0 = ct_0 t_{0S},$$

and from the fact that c is in the central of G we obtain the relation $t_{0S} = c_0$; that is

$$S^{-1} t_0 S = t_0 c_0.$$

But S and s transform the operations of G in the same way: therefore

$$\text{a)} \quad s^{-1} t_0 s = t_0 c_0.$$

Since $s^{-1}ts = tc$ it follows that $t^{-1}st = sc^{-1}$.

Now there exists an inner isomorphism which transforms the operations of G according to t. By an argument parallel to that employed above we can show that

$$\text{b)} \quad t^{-1} s_0 t = s_0 c_0^{-1}.$$

Relations a) and b) will be used to simplify the equation

$$\text{c)} \quad s_0^{-1} s^{-1} tt_0 ss_0 = tt_0 cc_0.$$

Now

$$s_0^{-1} s^{-1} tt_0 ss_0 = s_0^{-1} \cdot s^{-1} ts \cdot s^{-1} t_0 s \cdot s_0 = s_0^{-1} tct_0 c_0 s_0.$$

Substituting in c),

$$s_0^{-1} tct_0 c_0 s_0 = tt_0 cc_0.$$

Both c and c_0 are invariant under G; therefore

$$s_0^{-1} t t_0 s_0 = t t_0.$$

From the theorem of § 1 we know that s_0 must be commutative with t, since it is commutative with its conjugate $t t_0$ under T. Since $t^{-1} s_0 t = s_0 c_0^{-1}$, from relation b), it follows that c_0 must be the identity. In other words, *every commutator of G is characteristic.*

From the fact that c is characteristic, relations a) and b) reduce to $s^{-1} t_0 s = t_0$ and $t^{-1} s_0 t = s_0$ respectively. If s and t are commutative, then, by the theorem of § 1, s must be commutative with every multiplier of itself and of t. The obvious conclusion is stated in the corollary: *every multiplier of G is invariant under G.**

3. **Non-divisibility of G.** A group which is the direct product of two groups is said to be a *divisible group.* We shall now show that G cannot be divisible with one of its factors abelian. Let G be the direct product of an abelian group A and a non-abelian group N. It is clear that A must be cyclic; for a non-cyclic abelian group always admits isomorphisms which are not invariant under its group of isomorphisms. We assume, then, that A is generated by t of order p^a. Let s, of order p^n, be an operation of highest order in N.

If $a \leqq n$, we can construct an automorphism T_1 of G which will transform s into st; we can also find a second automorphism T_2 for which

a) $T_2^{-1} s T_2 = s$ and $T_2^{-1} t T_2 = ts'$, where $s' = t^p$, if $a > 1$, or, if t is of order p, where s' is an invariant operation of order p from N.

If $a > n$, we can construct an automorphism T_1' which will transform s into $st^{p^{a-n}}$; we can also establish an automorphism T_2' for which

b) $T_2'^{-1} s T_2' = s$ and $T_2'^{-1} t T_2' = ts'$, where $s' = t^p$, if $n > 1$. If $n = 1$, then, from the assumption that s is an operation of highest order in N, it follows that every operation in N is also of order p. Since a group in which every operation is of order 2 must be abelian, p must be an odd prime for the reason that N is non-abelian. If $p > 2$, then we may select t as the multiplier s' of t under T_2'.

In every case, then, we can find an automorphism (T_2 or T_2') which is commutative with s but not with all the conjugates of s under the group of isomorphisms of G. From this contradiction of the result expressed

* In view of the fact that the group generated by the commutators of G is a subgroup of the group generated by the multipliers of G, it is not surprising that the multipliers are subject to a weaker restriction. In the example given by G. A. Miller (loc. cit.) there is a multiplier of order 4 which is not characteristic.

in the theorem of Art. 1 we have proved that *the group G cannot be a divisible group with an abelian group as one of its factors.*

4. Generating operations of G. It is a well known fact that the number of independent generators of a group of order p^m is an invariant of the group.* In the case of an abelian group of order p^m its independent generators may be so selected that the group is the direct product of the cyclic subgroups generated by each of its independent generators. A direct analogy does not exist in the case of the group G which we are considering. However, we can show that, if p is an odd prime, then any set of independent generators $(s_1, s_2, \cdots, s_n)$ of the group G can be replaced by a set of independent generators $(s_1', s_2', \cdots, s_n')$ which have the property that no two of them shall have any powers in common except the identity.

Let us suppose that the orders of $s_1, s_2, \cdots, s_n$ are $p^{\alpha_1}, p^{\alpha_2}, \cdots, p^{\alpha_n}$ respectively, where $\alpha_1 \geqq \alpha_2 \geqq \cdots \geqq \alpha_n$. We shall assume that p^{β_2} is the lowest power of p for which $s_2^{p^{\beta_2}}, s_3^{p^{\beta_2}}, \cdots, s_n^{p^{\beta_2}}$ are each in $\{s_1\}$ while the p^{β_2-1} power of at least one of the generators, say s_i, is not in $\{s_1\}$. Since $\alpha_i \leqq \alpha_1$ there must be a power of s_1, say $s_1^{a_1}$ whose p^{β_2} power is the inverse of $s_i^{p^{\beta_2}}$. Then the order of $s_i s_1^{a_1}$ is p^{β_2}. For

$$(s_i s_1^{a_1})^{p^{\beta_2}} = s_i s_1^{a_1} s_i^{-1} \cdot s_i^2 s_1^{a_1} s_i^{-2} \cdot \cdots \cdot s_i^{p^{\beta_2}} s_1^{a_1} s_i^{-p^{\beta_2}} \cdot s_i^{p^{\beta_2}}.$$

If $s_i s_1^{a_1} s_i^{-1} = s_1^{a_1} c$, then

$$(s_i s_1^{a_1})^{p^{\beta_2}} = s_1^{a_1} c s_1^{a_1} c^2 \cdots s_1^{a_1} c^{p^{\beta_2}} s_i^{p^{\beta_2}}.$$

Since the commutators of G are invariant

$$(s_i s_1^{a_1})^{p^{\beta_2}} = s_1^{a_1 p^{\beta_2}} s_i^{p^{\beta_2}} c^{\frac{1}{2} p^{\beta_2}(p^{\beta_2}+1)}.$$

Now the order of c cannot exceed p^{β_2}. If $p > 2$, then $c^{\frac{1}{2} p^{\beta_2}(p^{\beta_2}+1)} = 1$. But $(s_1^{a_1})^{p^{\beta_2}} = s_i^{-p^{\beta_2}}$. Therefore

$$(s_i s_1^{a_1})^{p^{\beta_2}} = s_i^{-p^{\beta_2}} s_i^{p^{\beta_2}} = 1.$$

If we denote the product $s_i s_1^{a_1}$ by s_2', we see that the generator s_i may be replaced by the operation s_2' of order p^{β_2}, which has no powers in common with s_1, except the identity. This last statement is obvious from the fact that one of the constituents of s_2' is in $\{s_1\}$ while the other is not.

Suppose now that p^{β_3} is the lowest power of p for which the p^{β_3} power of each of the $(n-2)$ generators obtained by excluding s_1 and s_2' is in

* G. A. Miller: Trans. Amer. Math. Soc., vol. 16 (1915), p. 21.

either $\{s_1\}$ or $\{s_2'\}$, while the p^{β_2-1} power of at least one of them, say s_j, is neither in $\{s_1\}$ nor in $\{s_2'\}$. We may note that if $s_j^{p^{\beta_2}}$ is in both $\{s_1\}$ and $\{s_2'\}$ then $s_j^{p^{\beta_2}} = 1$; otherwise, $\{s_1\}$ and $\{s_2'\}$ would contain operators in common other that the identity. If $s_j^{p^{\beta_2}}$ is in $\{s_1\}$ and is not the identity, then the argument employed above may be used without modification to obtain an operation s_3' of order p^{β_2} which has no powers in common with s_1 except the identity; and it is further clear that the operation s_3' which is of the form $s_j\, s_1^b$ can have no powers in common with s_2' except the identity. If $s_j^{p^{\beta_2}}$ is in $\{s_2'\}$ and is not the identity, we might reasonably ask whether it would always be possible to find a power of s_2' whose p^{β_2} power would be the inverse of $s_j^{p^{\beta_2}}$. It is obviously possible to do so if the order of s_j does not exceed the order of s_2'. And if $\alpha_j > \beta_2$, then, remembering that $s_j^{p^{\beta_2}}$ is in $\{s_2'\}$ and is not the identity, it follows that $s_j^{p^{\alpha_j}}$ is necessarily the first power of s_j which is in $\{s_1\}$, a result which contradicts our assumption that the p^{β_2} power of every generator is in $\{s_1\}$. Regardless of whether $s_j^{p^{\beta_2}}$ is in $\{s_1\}$ or in $\{s_2'\}$, we can always find an operation s_3' of order p^{β_2} which has no powers in common with either s_1 or s_2' except the identity; and this operation s_3' can be used as a generator in place of s_j.

By a repetition of this argument it can be proved that *if* $p > 2$, *the generating operations of* G *may be selected in such a manner that they will have no powers in common except the identity.*

Let us see what statement can be made regarding the selection of the generators of G when $p = 2$. We remember that in the demonstration above it was necessary to show that $c^{\frac{1}{2}p^{\beta_2}(p^{\beta_2}+1)} = 1$, where c is a power of the commutator of s_1 and s_2. The proof depended on the fact that if p is an odd prime, then $p^{\beta_2}+1$ is divisible by 2. When $p = 2$ this is no longer the case; consequently it is possible that $c^{\frac{1}{2}(2^{\beta})(2^{\beta}+1)}$ is not the identity.

Let $(s_1, s_2, \cdots, s_n)$ be a set of independent generating operations of G of orders $2^{a_1}, 2^{a_2}, \cdots, 2^{a_n}$ respectively, where $a_1 \geqq a_2 \geqq \cdots \geqq a_n$. Let 2^{β} be the lowest exponent for which $s_2^{2^{\beta}}, s_3^{2^{\beta}}, \cdots, s_n^{2^{\beta}}$ are all in $\{s_1\}$, while the $2^{\beta-1}$ power of one of the set $(s_2, \cdots, s_n)$, say $s_2^{2^{\beta-1}}$, is not in $\{s_1\}$. Then there must be some power of s_1, say $s_1^{p^{2\alpha}}$ which is the inverse of $s_2^{2^{\beta}}$. (Since the following argument does not depend upon the value of k, we shall take k equal to unity.) It is evident that $a_1 - a_2 = \alpha - \beta$.

We shall now determine the order of $s_2\, s_1^{2^{\alpha-\beta}}$, assuming that $s_2\, s_1\, s_2^{-1} = s_1\, c_1$ where the order of c obviously cannot exceed 2^{β}. By raising $s_2\, s_1^{2^{\alpha-\beta}}$ to its m-th power, we obtain the following relations:

$$(s_2\, s_1^{2^{\alpha-\beta}})^m = s_2\, s_1^{2^{\alpha-\beta}}\, s_2^{-1} \cdot s_2\, s_1^{2^{\alpha-\beta}}\, s_2^{-2} \cdot \cdots \cdot s_2^m\, s_1^{2^{\alpha-\beta}}\, s_2^{-m} \cdot s_2^m$$
$$= s_1^{m2^{\alpha-\beta}}\, s_2^m \cdot c^{2^{\alpha-\beta}(m)(m+1)/2}.$$

If we set m equal to 2^β we see that $(s_2\, s_1^{2^{\alpha-\beta}})$ reduces to $c^{2^{\alpha-1}(2^\beta+1)}$, since $s_1^{2^\alpha} = s_1^{-2^\beta}$. If $\alpha > \beta$, then $c^{2^{\alpha-1}(2^\beta+1)} = 1$ and the order of $s_2\, s_1^{2^{\alpha-\beta}}$ is 2^α; hence we could replace s_2 in $(s_1, s_2, \cdots, s_n)$ by $s_2' = s_2\, s_1^{2^{\alpha-\beta}}$, where s_1 and s_2' obviously have only the identity in common. If $\alpha = \beta$ and the order of c is less than 2^β, the conclusion of the preceding sentence still holds.

If, now, $\alpha = \beta$ and the order of c is equal to 2^β, it is clear that the order of $s_2\, s_1^{2^{\alpha-\beta}} = s_2\, s_1$ equals $2^{\beta+1}$, since s_2^α is the inverse of s_1^α. Even here we can take $s_2\, s_1$ for the operation s_2' which is to replace s_2, unless s_1 and $s_2\, s_1$ have in common an operation of order 2, namely $c^{2^{\beta-1}(2^\beta+1)}$. If s_1 and $s_2\, s_1$ have in common this operation of order 2 and if $\beta + 1 < a$, where p^a is the common order of s_1 and s_2, then we can choose for s_2' the operation $s_2\, s_1^{1+2^{\alpha-\beta-1}}$ of order 2^β which has no powers in common with s_1 except the identity. That the order of $s_2\, s_1^{1+2^{\alpha-\beta-1}}$ is 2^β may be seen from the following equations:

$$(s_2\, s_1^{1+2^{\alpha-\beta-1}})^{2^\beta} = s_1^{2^\beta+2^{\alpha-1}} \cdot s_2^{2^\beta} \cdot c^{(1+2^{\alpha-\beta-1})2^{\beta-1}(2^\beta+1)}$$
$$= s_1^{2^{\alpha-1}} \cdot c^{2^{\beta-1}} = 1.$$

(We remember that $a - \beta > 1$; hence $c^{2^{a-2}} = 1$. Also, $s_1^{2^{\alpha-1}}$ and $c^{2^{\beta-1}}$ are the same operation of order 2.)

To recapitulate: Let $(s_1, s_2, \cdots, s_n)$ be a set of independent generating operations of G of orders $2^{a_1}, 2^{a_2}, \cdots, 2^{a_n}$ respectively, where $a_1 \geqq a_2 \geqq \cdots \geqq a_n$; if 2^β be the lowest exponent for which $s_2^{2^\beta}, \cdots, s_n^{2^\beta}$ are each in $\{s_1\}$ while the $2^{\beta-1}$ power of one of the set $(s_2, \cdots, s_n)$ say $s_j^{2^{\beta-1}}$, is not in $\{s_1\}$; then, either (1) it is possible to find an operation s_2' of order 2^β which has no operations in common with s_1 except the identity and which can replace s_j in the given set of generating operations, or (2) s_1 and s_j are of the same order 2^a, their commutator $c = s_1^{-1}\, s_j\, s_1\, s_j^{-1}$ is of order 2^{a-1}, the operations s_1, s_j have in common only the subgroup generated by the commutator $c^{2^{a-2}}$ of order 2.

From this point on the discussion follows the lines developed at the corresponding stage in that case $p > 2$, with the necessary modifications suggested by the results in the preceding paragraph. Finally we arrive at the following theorem, which logically falls into two sections according to the value of p.

If $p>2$, a set of independent generating operations of G can be found which shall possess the property that no two of them shall have in common any powers except the identity.

If $p=2$, a set of independent generating operations of G can be found which shall possess the property that any two of them have in common at most a commutator of G of order 2.

A set of generating operations for which this property holds shall be referred to as a *restricted set* of generating operations. In what follows it will be assumed that when reference is made to a set of independent generating operations, it is to a set selected in this manner.

5. **The order of I.** Since the group of inner isomorphisms of G is simply isomorphic with the central quotient group of G, it is clear that its order is a power of p. In this section we shall show that the order of every isomorphism of G is likewise a power of p. We shall employ a method of attack which is suggested by the theorem* that *every automorphism of an abelian group A may be obtained* (1) *by making A isomorphic with one of its subgroups or with itself in such a manner that no operation besides the identity corresponds to its inverse, and* (2) *by making each operation of A correspond to itself multiplied by the operation which corresponds to it in this isomorphism.*

Since every multiplier of G is invariant in G (cf. § 2), with every automorphism of G there is associated an isomorphism of G with a subgroup K of the central of G. Let s represent any operation of a set of independent generators of G and assume that s corresponds to sc under the automorphism T. If we denote by H the subgroup of G which is composed of those operations which are commutative with T and by K the subgroup of the central of G with which G is isomorphic under the automorphism T, then the correspondence of the operations of G under T may be exhibited by the following scheme:

$$\text{A)} \quad \begin{aligned} H &\sim H\cdot 1,\\ Ht_2 &\sim Ht_2\cdot c_2,\\ &\cdots\cdots\\ Hs &\sim Hs\cdot c \end{aligned}$$

where $1, c_2, \cdots, c$ are operations of K.

If T is commutative with c, it is evident that the exponent of the first power of T which is commutative with s is a power of p. Let us suppose, then, that T and c are not commutative; that is, we assume the relations

* Miller, Blichfeldt, and Dickson: Finite Groups, p. 103.

$$T^{-1}sT = sc, \qquad T^{-1}cT = c\gamma \qquad (\gamma \neq 1).$$

By hypothesis, s is a generating operation of G; consequently we may take c as a generating operation of the abelian subgroup K. From the automorphism T we may now construct an automorphism T' by replacing (vide scheme A) in K the multiplier c by the identity, which will have the effect of enlarging H by the adjunction of s and of other operations which under T are multiplied by powers of c. Since T' is commutative with s, it must not be commutative with c, by the theorem of Art. 1. This demands that γ be a power of c. Then, since $\gamma = c^\beta$, it follows that the cosets Hc and Hs^β associated with the automorphism T must be identical (vide scheme B). This implies that $c = s^\beta r$, where r is an operation of H.

B)
$$H \sim H \cdot 1,$$
$$Hc = Hs \sim Hc \cdot c^\beta = Hc \cdot \gamma,$$
$$\cdots\cdots\cdots\cdots$$
$$Hs \sim Hs \cdot c.$$

If γ and c are of the same order, then c and s must likewise have a common order; if, in addition, s is a non-invariant operation of G, then c, which equals $s^\beta r$, must also be a non-invariant (since s^β does not occur in H). This, however, contradicts the theorem that the multipliers of G are invariant. If s is an invariant in G and c is of the same order as s, then, since the operations of the commutator subgroup of G are characteristic, the first power of s which is in the commutator subgroup must be the identity; the group G would then be divisible with $\{s\}$ for one of its factors, which is contrary to the result established in § 3. We see, therefore, that γ is of lower order than c; as a consequence, from the relation $c = s^\beta r$, it follows that c has as a constituent a power of s of lower order than the order of s.

We may note that the operation r of H which occurs in the equation $c = s^\beta r$ must be a characteristic operation of G. Suppose that there exists an automorphism T_1 of G in which r corresponds to rr', $(r' \neq 1)$. It may be assumed that s and r have no powers in common except the identity; consequently, from the automorphism T_1 we can construct a second automorphism T_2 in which r corresponds to rr' and s corresponds to itself. But T_2 is commutative with s and not with c. Hence r must be a characteristic operation of G.

To recapitulate: If $T^{-1}sT = sc$ and $T^{-1}cT = c$, then the exponent of the first power of T which is commutative with s is a power of p; if $T^{-1}sT = sc$ and $T^{-1}cT = c\gamma$, then c is of lower order than s and γ

is of lower order than c. Combining these two results, we may assert that *the exponent of the first power of any automorphism of G which is commutative with a generating operation of G is a power of p.*

It is now easy to show that the order of every automorphism of G is a power of p. Since the commutators of G are invariant, any operation t of G may be written in the form

$$t = s_1^{x_1} s_2^{x_2} \cdots s_n^{x_n} s_0$$

where $s_1, s_2, \cdots, s_n$ are a set of generating operations and s_0 is in the commutator subgroup of G. (It is understood that $s_1, s_2, \cdots, s_n$ constitute a restricted set of independent generators of it [§ 4]). Let T be any automorphism of G and let $T^{p^{\alpha_1}}, T^{p^{\alpha_2}}, \cdots, T^{p^{\alpha_n}}$ be the first powers of T which are commutative with $s_1^{x_1}, s_2^{x_2}, \cdots, s_n^{x_n}$ respectively. It is easily seen that t and $T^{p^{\alpha_i}}$ are commutative, where p^{α_i} is the largest of the exponents $p^{\alpha_1}, p^{\alpha_2}, \cdots, p^{\alpha_n}$. Therefore the exponent of the first power of T which is commutative with t is a divisor of p^{α_i} and, consequently, the order of T is a power of p. In other words, *the order of the group of isomorphisms of G is a power of p.*

From the discussion above it is obvious that the order of any automorphism of G it at most equal to the order of an operation of highest order in the subgroup K associated with T, (vide scheme B). That is, *the order of every automorphism of G is a divisor of the order of an operation of highest order in the central of G.*

6. **Restrictions on the multipliers.** Let s be any one of a set of restricted generating operations of G. In § 5 it was proved that if $T^{-1} s T = s c$ and $T^{-1} c T = c\gamma$, where $\gamma \neq 1$, then $\gamma = c^{k^2 p^\beta}$ and $c = s^{k p^\beta} r$, where r is a characteristic operation of G. In the proof of the relation $c = s^{k p^\beta} r$ no use was made of the assumption $\gamma \neq 1$. Therefore we have in every case the result: *Every multiplier of a generating operation* s of G is the product of a characteristic operation of G and a power of s which is invariant under G.*

In § 5 it was also proved that the order of the group of isomorphisms I of G is a power of p; consequently, the number of conjugates of s under I is a power of p, say p^λ. Since every operation in I which is commutative with s is commutative with every conjugate of s under I, it follows that I, regarded as a substitution group on the p^m letters of G, must contain a subgroup J of order p^λ having a transitive constituent of degree p^λ under which s is transformed into all of its conjugates under I. Let us denote

* Vide p. 515.

the operations of J by 1, $T_2, \cdots, T_{p^\lambda}$ and let us assume the following relations:

$$T_i^{-1} s T_i = s c_i, \quad T_j^{-1} s T_j = s c_j, \quad T_i^{-1} c_j T_i = c_j c_{ij},$$
$$T_j^{-1} c_i T_j = c_i c_{ji}.$$

By the argument which was used in § 1 it can be proved that $c_i c_j c_{ij} = c_j c_i c_{ji}$.

Since the multipliers of G are in the central of G (cf. § 2) it follows that $c_{ij} = c_{ji}$. Suppose that $T_\alpha^{-1} s T_\alpha = s c_\alpha$, where c_α is a characteristic operation of G; then $c_{\alpha i} = c_{i\alpha} = 1$, which implies that T_α is commutative with every multiplier of s under I.

If $T_\alpha^{-1} s T_\alpha = s c_\alpha$ and $T_\beta^{-1} s T_\beta = s c_\beta$ while $c_{\alpha\beta} \neq 1$, then c_β must contain as a constituent a power of s, since every multiplier of s is of the form $s^{kp^\beta} r$ where r is a characteristic operation of G. Let $c_\beta = s^{kp^\beta} r$. Now

$$T_\alpha^{-1} T_\beta^{-1} s T_\beta T_\alpha = T_\alpha^{-1} s s^{kp^\beta} r T_\alpha = s c_\alpha s^{kp^\beta} c_\alpha^{kp^\beta} r,$$

$$T_\beta^{-1} T_\alpha^{-1} s T_\alpha T_\beta = T_\beta^{-1} s c_\alpha T_\beta \quad = s s^{kp^\beta} r c_\alpha c_{\beta\alpha}.$$

Since $T_\alpha T_\beta = T_\beta T_\alpha$,

$$c_\alpha s^{kp^\beta} c_\alpha^{kp^\beta} r = s^{kp^\beta} r c_\alpha c_{\beta\alpha},$$

whence $c_{\beta\alpha} = c_\alpha^{kp^\beta}$. From the relation $c_{\beta\alpha} = c_{\alpha\beta}$ and from the fact that c_β is not characteristic [$c_{\alpha\beta} \neq 1$, by assumption] one sees that c_α must involve a non-characteristic power of s; i. e.,

$$c_\alpha = s^{k' p^\alpha} r'$$

where r' is a characteristic operation. It is obvious that J contains an operation T_r which transforms s into sr. Then

$$T_r^{-1} c_\alpha T_r = T_r^{-1} s^{k' p^\alpha} r' \cdot T_r = s^{k' p^\alpha} r' \cdot r^{k' p^\alpha}.$$

In the preceding paragraph it was proved that an automorphism which transforms s into itself multiplied by a characteristic operation of G must be commutative with every multiplier of s under I. Therefore $r^{k' p^\alpha} = 1$. Similarly it can be proved that $r^{kp^\beta} = 1$, $r'^{k' p^\alpha} = 1$, and $r'^{kp^\beta} = 1$. From these relations and the relations $c_\alpha = s^{k' p^\alpha} r'$, $c_\beta = s^{kp^\beta} r$, $c_{\beta\alpha} = c_\alpha^{kp^\beta}$, $c_{\alpha\beta} = c_{\beta\alpha}$, we see that $c_{\alpha\beta}$ and $c_{\beta\alpha}$ have the common value $s^{kk' p^{\alpha+\beta}}$. This result may be stated in the following form: *if the multipliers of a generating operation s are not all characteristic, then I must contain an operation T_α which transforms s into $s \cdot s^{kp^\alpha} r$ and transforms s^{kp^α} into itself multiplied by $s^{k^2 p^{2\alpha}}$, where k is prime to r, $\alpha > 0$, s^{kp^α} is in the central of G, $s^{p^{2\alpha}} \neq 1$, and r is a characteristic operation of G.*

It is clear that the multipliers $1, c_2, \cdots, c_{p^\lambda}$ associated with the operations $1, T_2, \cdots, T_{p^\lambda}$ of J are all distinct. One readily sees that, if s corresponds to sc_α under T_α and to sc_β under T_β, then J must contain an automorphism under which s corresponds to $sc_\alpha c_\beta$. We note that this automorphism is the product $T_\alpha T_\beta$ if, and only if, c_α and T_β are commutative. Since the product of two multipliers is itself a multiplier, we conclude that the multipliers of s form a subgroup Γ of order p^λ. Evidently the subgroup Γ contains the multipliers c_{ij}. It is clear that Γ and J are simply isomorphic if the multipliers of s are characteristic. Suppose that Γ contains a multiplier which is not characteristic. By the theorem of the preceding paragraph we know that J must contain an operation T_α for which

$$T_\alpha^{-1} s T_\alpha = s s^{kp^\alpha} r \text{ and } T_\alpha^{-1} s^{kp^\alpha} r T_\alpha = s^{kp^\alpha} r s^{k^2 p^{2\alpha}}$$

where $s^{k^2 p^{2\alpha}}$ is not the identity. It has been proved above that the order of every characteristic multiplier is a divisor of p^α if s^α is a multiplier of s which is not characteristic. We see, then, that the order of T_α is a divisor of the order of s^{kp^α}. Suppose that T_α is of order* p^b. Then†

$$T_\alpha^{-p^b} s T_\alpha^{p^b} = s^{(1+p^\alpha)p^b}.$$

Now

$$(1+p^\alpha)^{p^b} = 1 + p^{b+\alpha}\left[1 + \frac{p^b-1}{2} p^\alpha + \cdots + \frac{(p^b-1)!\, p^{(r-1)\alpha}}{r!\,(p^b-r)!} + \cdots\right].$$

If the order of T_α is less than the order of s^{p^α}, then the quantity Q in the brackets must be divisible by p. If p is an odd prime, then the coefficients of the powers of p contained in Q are all integers and Q is of the form $1 + kp$. If $p = 2$ and $\alpha > 1$, then $Q \equiv 1 \bmod 2$. If $\alpha = 1$, then $(1+2)^{2^b}$ is represented by the following expansion:

$$\begin{aligned}(1+2)^{2^b} = 1 + 2^{b+1}\Big[1 &+ (2^b-1) + \frac{2^b-1}{3}(2^{b-1}-1)2^2 \\ &+ \frac{2^b-1}{3}(2^{b-1}-1)(2^b-3)2 \\ &+ \frac{2^b-1}{3\cdot 5}(2^b-1)(2^b-3)(2^{b-2}-1)2^2 + \text{terms divisible by } 2^3\Big].\end{aligned}$$

* Since I is abelian and s is a generating operation of Q, it may be assumed that T_α is commutative with the remaining generating operations of the set to which s belongs. Then the order of T_α is the first power of T_α which is commutative with s.

† We note that $r^{p^\alpha} = 1$. Since the following discussion does not depend on the value of k, we shall, for convenience, take $k = 1$.

We see that in this case Q is even. If the order of s is greater than 4, then p^b, the order of T_α, divides one half the order of $s^{2^\alpha} = s^2$ and, of course, one-fourth the order of s. We see, therefore, that only in the exceptional case where $p = 2$ and $\alpha = 1$ can the order of T_α be less than the order of the multiplier of T_α and s. The general result may be stated in the following theorem: *if p is an odd prime, then the subgroup J of the group of isomorphisms I of G which is composed of those operations of I which transform a generating operation s of G into all its conjugates and are commutative with the remaining generating operations of G is simply isomorphic with the subgroup of the central of G which is generated by the multipliers of s. The same statement can be made when $p = 2$ if either of the following conditions is satisfied:* (1) *the order of s does not exceed* 4; (2) *there is no operation in J which transforms s into $s \cdot s^{2k}$ where $k \equiv 1 \bmod 2$. If there is an operation T in J which transforms s into $s \cdot s^{2k}$ and if the order of s is greater than* 4, *then the order of T divides one-fourth of the order of s.*

OSCILLATION THEOREMS FOR THE DIFFERENTIAL BOUNDARY VALUE PROBLEMS OF THE FOURTH ORDER.*

By S. A. Janczewsky.

Since Sturm the oscillation theorems for the linear differential equations the second order have been discussed by many authors. This is not the case, however, with the problems of higher orders, particularly with those of the fourth order. The literature of the question consists of but few papers, among which should be mentioned those by Davidoglou,† the Thesis by O. Haupt‡ and some general remarks by R. D. Carmichael.§ The purpose of the present note is to contribute to the oscillation theory of a class of Sturmian boundary value problems of the fourth order. The method used throughout the paper is the so called continuity method and the principal results have been enunciated briefly in our two notes in the Comptes Rendus.‖

1. A Sturmian system of the 4th order is defined as a self-adjoint boundary value problem consisting of a self-adjoint differential equation

$$L(y) \equiv [\varrho(x)y'']'' + [\psi(x) - \lambda\tau(x)]y = 0 \qquad (a \leqq x \leqq b) \tag{1}$$

together with a set of four linearly independent boundary conditions of which one pair bear to the end-point $x = a$ and another pair to the end-point $x = b$:

$$\begin{aligned} U_1(y) &\equiv a_{11}(\varrho y'')' + a_{12}(\varrho y'') + a_{13}y' + a_{14}y]_{x=a} = 0, \\ U_2(y) &\equiv a_{21}(\varrho y'')' + a_{22}(\varrho y'') + a_{23}y' + a_{24}y]_{x=a} = 0, \\ U_3(y) &\equiv b_{11}(\varrho y'')' + b_{12}(\varrho y'') + b_{13}y' + b_{14}y]_{x=b} = 0, \\ U_4(y) &\equiv b_{21}(\varrho y'')' + b_{22}(\varrho y'') + b_{23}y' + b_{24}y]_{x=b} = 0. \end{aligned} \tag{2}$$

* Received January 18, 1928.

† Sur l'équation des vibrations transversales des verges élastiques, Annales Sc. de l'École Norm. Sup. (3), 17 (1900), pp. 359–444, Étude du l'équation différentielle $(\theta y'')'' = K\varphi(x)y$, ibidem, (3), 22 (1905), pp. 539–565.

‡ Untersuchungen über Oszillationstheoreme, Würzburg, 1911.

§ Comparison theorems for homogeneous linear differential equations of general order, Ann. of Math. (2) 19 (1917–18), pp. 159–171; Boundary value and expansion problems: oscillation, comparison and expansion theorems, Amer. Journ. Math., 44 (1922), pp. 129–152.

‖ Sur quelques problèmes aux limites pour les équations différentielles linéaires ordinaires du quatrième ordre, C. R. 184 (1927), pp. 141–143; Les théorèmes d'oscillation des problémes reguliers de Sturm pour les équations différentielles linéaires ordinaires du quatrième ordre, ibidem, pp. 261–263.

Here $\varrho(x)$, $\psi(x)$ and $\tau(x)$ are given functions in x on (a, b) and a_i, b_i $(i = 1, 2, 3, 4)$ are given constants which satisfy the conditions

$$(3) \qquad \begin{vmatrix} a_{11} & a_{14} \\ a_{21} & a_{24} \end{vmatrix} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \quad \begin{vmatrix} b_{11} & b_{14} \\ b_{21} & b_{24} \end{vmatrix} = \begin{vmatrix} b_{12} & b_{13} \\ b_{22} & b_{23} \end{vmatrix}$$

which are necessary and sufficient that the system (2) be self-adjoint.* Denote by $(\alpha\beta\gamma\delta)$ the determinant of the matrix

$$(4) \qquad \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & 0 & 0 & 0 & b_{21} & b_{22} & b_{23} & b_{24} \end{pmatrix}$$

with the column numbers $\alpha, \beta, \gamma, \delta$ respectively. Among these determinants only 36 may happen to be different from zero. We order them as follows:

$$(5) \qquad \begin{gathered} (1256),\ (1257),\ (1356),\ (1258) = (1267),\ (1357),\ (1456) = (2356), \\ (1268),\ (1358) = (1367),\ (1457) = (2357),\ (2456),\ (1278),\ (1368), \\ (1458) = (1467) = (2358) = (2367),\ (2457),\ (3456),\ (1378), \\ (1468) = (2368),\ (2458) = (2367),\ (3457),\ (1478) = (2378), \\ (2468),\ (3458) = (3467),\ (2478),\ (3468),\ (3478). \end{gathered}$$

The term „principal determinant“ will be used to designate any member of the set (5), which may be different from zero while all the preceding terms are equal to zero. In view of the relations (3) it is readily seen that only the determinants of (5) which are not of one of the forms $(14\gamma\delta)$, $(23\gamma\delta)$, $(\alpha\beta 58)$, $(\alpha\beta 67)$ are principal ones. It is also readily proved that the transformation $x' = a + b - x$ transforms a system whose principal determinant is†

$$(1257),\ (1268),\ (1278),\ (1368),\ (1378),\ (2478)$$

into a system whose principal determinant is respectively:

$$(1356),\ (2456),\ (3456),\ (2457),\ (3457),\ (3468).$$

---

* This can be verified directly and also is readily derived from a general result of D. Jackson, Algebraic properties of self-adjoint systems, Trans. Amer. Math. Soc., 17 (1916), pp. 418–424 (423).

† More exactly: „whose first principal determinant of the set (5), which is different from zero, is . . .“.

We shall not consider as distinct the systems which can be transformed one into another by the transformation $x' = a + b - x$. Then an easy computation will show that any Sturmian system of the fourth order is equivalent to one of the ten canonical forms below, according as which is the first principal determinant $\neq 0$:

Let

$$
\begin{aligned}
&A_1(y) \equiv (\varrho y'')' - A_0 y' + A_1 y]_{x=a}; && B_1(y) \equiv (\varrho y'')' - B_0 y' + B_1 y]_{x=b}\\
&A_2(y) \equiv \varrho y'' + A_2 y' + A_0 y]_{x=a}; && B_2(y) \equiv \varrho y'' + B_2 y' + B_0 y]_{x=b}\\
(6)\quad &A_3(y) \equiv (\varrho y'')' + A_3 \varrho y'' + A_4 y]_{x=a}; && B_3(y) \equiv (\varrho y'')' + B_3 \varrho y'' + B_4 y]_{x=b}\\
&A_4(y) \equiv y' + A_3 y]_{x=a}; && B_4(y) \equiv y' + B_3 y]_{x=b}\\
&A_5(y) \equiv \varrho y'' + A_5 y']_{x=a}; && B_5(y) \equiv \varrho y'' + B_5 y']_{x=b}.
\end{aligned}
$$

Then the canonical forms in question are given by the table

1·st PD $\neq 0$	i (1256)	ii (1356)	iii (1357)	iv (2456)	v (2457)	vi (2468)	vii (3456)	viii (3457)	ix (3468)	x (3478)
Canonical form:	$A_1(y)$	$A_3(y)$	$A_3(y)$	$A_5(y)$	$A_5(y)$	$A_5(y)$	$y'(a)$	$y'(a)$	$y'(a)$	$y'(a)$
	$A_2(y)$	$A_4(y)$	$A_4(y)$	$y(a)$	$y(a)$	$y(a)$	$y(a)$	$y(a)$	$y(a)$	$y(a)$
$0 =$	$B_1(y)$	$B_1(y)$	$B_3(y)$	$B_1(y)$	$B_3(y)$	$B_5(y)$	$B_1(y)$	$B_3(y)$	$B_5(y)$	$y'(b)$
	$B_2(y)$	$B_2(y)$	$B_4(y)$	$B_2(y)$	$B_4(y)$	$y(b)$	$B_2(y)$	$B_4(y)$	$y(b)$	$y(b)$

2. In this paper we are concerned exclusively with the Sturmian systems in which the functions $\varrho''(x)$, $\psi(x)$, $\tau(x)$ are continuous on (a, b), $\varrho(x) > 0$, $\tau(x) \geqq 0$, and $\not\equiv 0$ on any subinterval of (a, b). This assumption will be made throughout the paper, without separate mentioning, when we speak of a Sturmian system.

A Sturmian system will be designated as *regular* if, in addition, the coefficients of the boundary conditions in a canonical form satisfy the conditions:

$$(7)\qquad A_0 \geqq 0,\quad B_0 \geqq 0,\quad B_2 \geqq 0,\quad B_3 \geqq 0,\quad A_2 \leqq 0,\quad A_3 \leqq 0.$$

A regular Sturmian system is called *completely regular*, if, besides the conditions (7), we have also:

$$
\begin{gathered}
A_0^2 \leqq |A_1 A_2|,\quad B_0^2 \leqq |B_1 B_2|,\\
(8)\qquad A_1 > 0,\quad A_5 \leqq 0,\quad B_5 \geqq 0,\quad B_1 < 0,\\
A_4 > 0\,[\geqq 0 \text{ if } A_3 \not\equiv 0],\quad B_4 < 0\,[\leqq 0 \text{ if } B_3 \not\equiv 0].
\end{gathered}
$$

It is readily seen that the boundary conditions of a completely regular system can be brought to the form

$$\text{(xi)}\qquad \begin{aligned} A_6(y) &\equiv \varrho y'' - C_0(\varrho y'')' - C_1 y']_{x=a} = 0,\\ A_7(y) &\equiv \quad y + C_2(\varrho y'')' - C_0 y']_{x=a} = 0,\\ B_6(y) &\equiv \varrho y'' + D_0(\varrho y'')' + D_1 y']_{x=b} = 0,\\ B_7(y) &\equiv \quad y - D_2(\varrho y'')' + D_0 y']_{x=b} = 0 \end{aligned}$$

where

$$C_0, C_1, C_2, D_0, D_1, D_2 \geqq 0 \tag{9}$$

and the cases where $C_1 = \infty$ or $D_1 = \infty$ are not excluded, in order to incorporate conditions of the types

$$y'(a) = 0 \ (A_3 = 0); \quad y'(b) = 0 \ (B_3 = 0).$$

3. The characteristic values of the system (1, 2) are defined as the values of the parameter λ for which the system admits of at least one solution $\not\equiv 0$ on (a, b). To each characteristic value there corresponds at least one fundamental function of (1, 2), that is a solution of (1, 2) which is $\not\equiv 0$ on (a, b).

The *index* of a characteristic value is defined as the number of linearly independent fundamental functions corresponding to this value. A characteristic value of index 1 is simple.

Two important identities should be quoted here:

The „Green's identity"

$$\begin{gathered} \int_a^b [z\,L(y) - y\,L(z)]\,dx = P(y, z)\Big]_a^b,\\ P(y, z) \equiv (\varrho y'')'z - \varrho y''z' + y'\varrho z'' - y(\varrho z'')'. \end{gathered} \tag{10}$$

The „Dirichlet's identity"

$$\begin{gathered} \int_a^b y\,L(y)\,dx = \int_a^b [\varrho y'^2 + \psi y^2 - \lambda \tau y^2]\,dx + Q(y)\Big]_a^b,\\ Q(y) \equiv y(\varrho y'')' - y'\varrho y''. \end{gathered} \tag{11}$$

Both identities hold true for any pair of functions $y(x)$, $z(x)$ whose derivatives up to the 4th order are continuous on (a, b), provided the functions $\varrho''(x)$, $\psi(x)$ and $\tau(x)$ are continuous on (a, b).

4. In the sequel we shall use some known properties of the solutions of linear boundary value problems of the fourth order:

1°. *No linear boundary value problem of the fourth order with* $\varrho(x) \neq 0$ *and continuous* $\varrho''(x)$, $\psi(x)$, $\tau(x)$ *can have a characteristic value whose index exceeds* 4. *The set of characteristic values has no finite limiting points.*

2°. *A solution of the differential equation*

$$L(y) = 0 \tag{12}$$

which is zero together with its derivatives up to the 3rd order at any point in (a, b), *is* $\equiv 0$ *on* (a, b).

COROLLARY. *A solution of* (12) *may have but a finite number of zeros on* (a, b).

3°. *If* $\tau(x) \geqq 0$ *and there exists no subinterval of* (a, b) *on which* $\tau(x) \equiv 0$, *the corresponding Sturmian system has infinitely many characteristic values.*

4°. *The characteristic values and the fundamental functions of a Sturmian system of* 3° *are real. The fundamental functions corresponding to the different characteristic values are orthogonal (to the base* $\tau(x)$), *that is, for any pair* $y_1(x)$, $y_2(x)$ *of fundamental functions in question,*

$$\int_a^b \tau(x)\, y_1(x)\, y_2(x)\, dx = 0.$$

5°. *If, in a Sturmian system,* $\psi(x) \geqq 0$ *and* $Q(y)\Big]_a^b \geqq 0$ *in virtue of the boundary conditions, then there exist no negative characteristic values.*

COROLLARY. *A completely regular Sturmian system with* $\psi(x) \geqq 0$ *has no negative characteristic values.*

6°. *The characteristic values of a Sturmian system with*

$$\varrho(x) \equiv 1, \quad \psi(x) \equiv 0, \quad \tau(x) \equiv \text{constant} = k > 0,$$

are bounded below, that is there exists a positive constant A_0 which depends on k and the coefficients of the boundary conditions of the problem, such that all the characteristic values are $> -A_0$. This constant A_0, as function of k and the coefficients A_i, B_i of the canonical forms of the boundary conditions, is uniformly bounded, provided k ranges over a finite interval $(\alpha, \beta)\,[0 < \alpha < \beta]$ and A_i; B_i remain bounded.

The statements 1°—2° are obvious. Without loss of generality we may assume that 0 is not a characteristic value of our Sturmian system.* Then there exists the Green's function $G(x, t)$ of the system

$$\begin{aligned} L_1(y) &\equiv [\varrho(x)y']' + \psi(x)\, y(x) = f(x), \\ U_i(y) &= 0 \qquad\qquad (i = 1, 2, 3, 4), \end{aligned} \tag{13}$$

$f(x)$ being an arbitrary continuous function. The solution of (13) is uniquely determined and given by

$$y(x) = \int_a^b G(x, t) f(t)\, dt.\dagger$$

* If 0 is a characteristic value but λ_0 is not, then the assumption of the text holds true for a system derived from the given one by substituting λ instead of $(\lambda - \lambda_0)$ and $\psi(x) - \lambda_0 \tau(x)$ instead of $\psi(x)$.

† Cfr. Westfall, Zur Theorie der Integralgleichungen, Thesis, Göttingen 1905; Birkhoff, Boundary value and expansion problems of ordinary linear differential equations, Trans. Amer. Math. Soc., 9 (1908), pp. 373–395.

Hence, on setting

$$K(x, t) \equiv G(x, t)\sqrt{\tau(x)\,\tau(t)}; \qquad u(x) \equiv y(x)\sqrt{\tau(x)} \tag{14}$$

we see that the system (13) is equivalent to the Fredholm integral equation

$$u(x) = \lambda \int_a^b K(x, t)\, u(t)\, dt. \tag{15}$$

The kernel $K(x, t)$ is continuous, symmetric and closed, that is, a relation of the form

$$\int_a^b K(x, t)\, h(t)\, dt \equiv 0$$

where $h(x)$ is a continuous function, implies $h(x) \equiv 0$.* The statements $3° - 4°$ now follow immediately from the known facts of the theory of integral equations.†

Statement $5°$ follows immediately from (11).‡

Finally, to prove the statement $6°$, we use the fact that the Sturmian systems mentioned there are regular in the sense of Birkhoff.§ Hence the characteristic values of the systems of $6°$ are given by

$$(2\nu\pi)^4/k \cdot \left(1 + \frac{E}{\nu}\right) \qquad (\nu = \pm 1, \pm 2, \cdots) \tag{16}$$

where E is bounded‖ under the hypotheses of $6°$.

5. LEMMA 1. *If λ be a characteristic value of a regular Sturmian system and $y(x)$ the corresponding fundamental function and if*

$$\psi(x) - \lambda\, r(x) < 0 \text{ on } (a, b) \tag{17}$$

then: $1°$. *The roots of $y(x)$ interior to (a, b) are simple.*

$2°$. *In the case of the boundary conditions* (i–iii) *or* (xi) *with $C_1 \neq \infty$ and $C_0 + C_2 \neq 0$ or with $C_1 = \infty$ and $C_2 \neq 0$ we have $y(a) \neq 0$.*

$3°$. *In the case of the boundary conditions* (iv–vi) *or* (xi) *with $C_1 \neq \infty$ and $C_0 = C_2 = 0$ we have $y'(a) \neq 0$.*

---

* For, on setting $z(x) \equiv \int_a^b G(x, t)\, h(t)\, dt$, we have $L_1(z) = h(x)$.

† See i. e. Lalesco, Introduction a la théorie des équations intégrales, Paris, 1912, 64-73. The proof of the text was given by Bounitzky, Concerning the solution of the ordinary linear differential equations under given boundary conditions, Odessa, 1916 (In Russian).

‡ Tamarkin, Some general problems of the theory of ordinary linear differential equations, Petrograd, 1917 (In Russian), pp. 137–139.

§ Loc. cit. footnote †, p. 525; the proof of the regularity of the systems in question has been given by Tamarkin, Sur quelques points de la théorie des équations différentielles linéaires ordinaires et sur la generalisation de la série de Fourier, Rend. Circ. Mat., Palermo, 34 (1912), pp. 345–382 (359).

‖ See footnote †, p. 525, loc. cit. pp. 382–387.

4°. *In the case of the boundary conditions* (vii–x) *we have* $y'(a) \neq 0$.

5°. *In the case of the boundary conditions* (i–v, vii, viii) *or* (xi) *with* $D_1 \neq \infty$ *and* $D_0 + D_2 \neq 0$ *or with* $D_1 = \infty$ *and* $D_2 \neq 0$ *we have* $y(b) \neq 0$.

6°. *In the case of the boundary conditions* (vi, ix) *or* (xi) *with* $D_1 \neq \infty$ *and* $D_0 = D_2 = 0$ *we have* $y'(b) \neq 0$.

7°. *In the case of the boundary conditions* (x) *we have* $y''(b) \neq 0$.

To prove 1°, suppose that

$$y(x_0) = y'(x_0) = 0 \qquad (a < x_0 < b).$$

Without loss of generality we may confine our discussion to the three cases as follows:

$$(\alpha) \qquad \varrho\, y''\big]_{x=x_0} > 0, \quad (\varrho\, y'')'\big]_{x=x_0} \geqq 0;$$

$$(\beta) \qquad \varrho\, y''\big]_{x=x_0} = 0, \quad (\varrho\, y'')'\big]_{x=x_0} > 0;$$

$$(\gamma) \qquad \varrho\, y''\big]_{x=x_0} = 0, \quad (\varrho\, y'')'\big]_{x=x_0} < 0.$$

We first observe that, if at any point $x = x_1$, we have

$$y,\ y',\ \varrho\, y'',\ (\varrho\, y'')'\big]_{x=x_1} > 0,$$

then it must be

$$y(x),\ y'(x),\ \varrho\, y'',\ (\varrho\, y'')' > 0 \text{ on } x_1 \leqq x \leqq b,$$

since

$$(\varrho\, y'')''\big]_{x=x_1} = -[\psi(x_1) - \lambda\, \tau(x_1)]\, y(x_1) > 0$$

and all the functions y, y', $\varrho\, y''$, $(\varrho\, y'')'$ being positive and increasing at $x = x_1$ will keep this property on the whole of (x_1, b). Now, in the case (α), y' is increasing at $x = x_0$, hence it becomes positive for $x = x_0 + h$ where $h > 0$ is sufficiently small, and the same will be true of y, $\varrho\, y''$, $(\varrho\, y'')'$. On setting $x_1 = x_0 + h$ in our remark above, we see at once that

$$(18) \qquad y(b),\quad y'(b),\quad \varrho\, y''\big]_{x=b},\quad (\varrho\, y'')'\big]_{x=b} > 0.$$

The same conclusion is readily obtained in the case (β).

As to the case (γ), it can be reduced to (α) by setting $x' = a + b - x$. Our conclusion in this case will be then that

$$(19) \qquad y(a) > 0,\quad y'(a) < 0,\quad \varrho\, y''\big]_{x=a} > 0,\quad (\varrho\, y'')'\big]_{x=a} < 0.$$

Since the inequalities (18), (19) are not compatible with any of the boundary conditions of a regular type, the statement 1° is proved. To prove 2°, take for instance the boundary conditions (i). If $y(a) = 0$ then conditions

$$(\varrho\, y'')'\big]_{x=a} = A_0\, y'(a),\quad \varrho\, y''\big]_{x=a} = -A_2\, y'(a),\quad A_0 \geqq 0,\quad A_2 \leqq 0.$$

in view of 3,2° show that $y'(a) \neq 0$, We may assume without loss of generality that $y'(a) > 0$. Hence if $x_1 = a + h$ ($h > 0$ and sufficiently small)

$$y(x_1), \quad y'(x_1), \quad \varrho y''\big]_{x=x_1}, \quad (\varrho y'')'\big]_{x=x_1} > 0$$

and also

$$y(b), \quad y'(b), \quad \varrho y''\big]_{x=b}, \quad (\varrho y'')'\big]_{x=b} > 0$$

which is not possible. The other cases of Lemma 1 can be treated in an analogous fashion.

COROLLARY. *All the characteristic values of a regular Sturmian system, for which* $\psi(x) - \lambda\tau(x) < 0$ *on* (a, b) *are simple.*

For, if λ is a characteristic value whose index exceeds 1, there exist at least two linearly independent fundamental functions corresponding to λ. If we denote them by $y_1(x)$ and $y_2(x)$ respectively, then, on choosing the constants α, β in a suitable way, we can construct a fundamental function

$$y(x) = \alpha y_1(x) + \beta y_2(x)$$

for which $y^{(k)}(a) = 0$,* where k is one of the numbers 0, 1, 2 taken at pleasure. This result, however, is non-compatible with Lemma 1.

Remark. Lemma 1 and Corollary hold true in the case $\psi(x) \equiv 0$, provided the characteristic value λ is > 0 and $\tau(x) \geqq 0$ can not be identically zero on any subinterval of (a, b).

6. Denote by

$$\lambda_1, \lambda_2, \cdots, \lambda_\nu, \cdots \quad (0 \leqq |\lambda_1| \leqq |\lambda_2| \leqq \cdots \leqq |\lambda_\nu| \leqq \cdots)$$

the set of the characteristic values of a given Sturmian system, where each characteristic value is repeated as many times as is indicated by its index.

If the coefficients of the operators $L(y)$ and $U_i(y)$ depend on a parameter μ, then the characteristic values and the fundamental functions become functions of μ. We assume that μ ranges over a finite closed interval I ($\mu_1 \leqq \mu \leqq \mu_2$). The characteristic values and the fundamental functions will be designated now by

$$\{\lambda_\nu(\mu)\} \quad \text{and} \quad \{\varphi_\nu(x, \mu)\}$$

respectively. The set $\{\varphi_\nu(x, \mu)\}$ is assumed to be orthogonal and normalized (to the base $\tau(x, \mu)$):

$$\int_a^b \tau(x, \mu)\, \varphi_i(x, \mu)\, \varphi_k(x, \mu)\, dx = \begin{cases} 0 \text{ if } k \neq i, \\ 1 \text{ if } k = i. \end{cases}$$

* or $y^{(k)}(b) = 0$.

In addition we assume that

1°. The functions $\varrho(x, \mu)$, $\varrho'(x, \mu)$, $\varrho''(x, \mu)$, $\psi(x, \mu)$, $\tau(x, \mu)$ are continuous in x and μ on $[(a, b), I]$.

2°. There exists a constant ϱ_0 which does not depend on μ and such that $\varrho(x, \mu) > \varrho_0 > 0$ on $[(a, b), I]$.

3°. For a fixed μ

$$\tau(x, \mu) \geqq 0$$

and there is no subinterval of (a, b) on which $\tau(x, \mu) \equiv 0$.

4°. When μ ranges over I, the type of the boundary conditions in the canonical form (i–x) remains the same and the corresponding coefficients are bounded and continuous in μ. The same is assumed in the case of the boundary conditions (xi).*

The set of all the conditions above we shall designate briefly by (*).

7. LEMMA 2. *Under the conditions* (*) *the characteristic values* $\{\lambda_\nu(\mu)\}$, *so far as they remain finite, are real and continuous in* μ. *If several characteristic values* $\lambda^{(1)}(\mu)$, $\lambda^{(2)}(\mu), \ldots$ *approach a common limit* λ_0 *as* $\mu \to \mu_0$, *then* λ_0 *is a characteristic value of the problem for* $\mu = \mu_0$ *and the index of* λ_0 *is equal to the sum of the indices of* $\lambda^{(1)}(\mu)$, $\lambda^{(2)}(\mu), \ldots$.

Let $\varphi(x, \mu)$ *be a (normalized) fundamental function which corresponds to one of* $\lambda^{(1)}(\mu), \ldots$. *Given a positive arbitrarily small* ε, *it is always possible to find a (normalized) fundamental function* $\varphi_0(x)$ *which corresponds to* λ_0 *and a positive number* $\delta(\varepsilon)$ *such that the difference*

$$\varphi(x, \mu) - \varphi_0(x)$$

and its derivatives up to the fourth order (with respect to x*) be* $< \varepsilon$ *in absolute value for* $a \leqq x \leqq b$, *provided* μ *is fixed and* $|\mu - \mu_0| < \delta(\varepsilon)$.

Remark. If λ_0 is not a simple characteristic value, Lemma 2 does not imply the continuity in μ of $\varphi(x, \mu)$, since the choice of $\varphi_0(x)$ may depend on μ. However, in the case where λ_0 is a simple characteristic value, the same is true of $\lambda^{(1)}(\mu)$ for μ sufficiently near to μ_0; the functions $\varphi(x, \mu)$ and $\varphi_0(x)$ then are uniquely determined and

$$\varphi(x, \mu) \to \varphi_0(x) \text{ as } \mu \to \mu_0,$$

so that $\varphi(x, \mu)$ is continuous in μ at $\mu = \mu_0$.

To prove Lemma 2, let μ_0 be any point of I. We may assume without loss of generality that 0 is not a characteristic value of the problem for $\mu = \mu_0$. The same will hold true, then, for all μ in an interval

* If one or both of the coefficients C_1, D_1 become ∞, it is assumed to remain ∞ as μ changes, so that the corresponding boundary condition always has the form $y'(a) = 0$ or $y'(b) = 0$. The following discussion can be applied as well to the case where the operators of our problem depend on several parameters, provided all the conditions above are satisfied.

$$I_{\mu_0} : \mu_0 - h \leqq \mu \leqq \mu_0 + h^*$$

where h is sufficiently small. Then the Green's function $G(x, t, \mu)$ of the system

$$[\varrho(x, \mu) y'']'' + \psi(x, \mu) y = f(x)$$
$$U_i(y, \mu) = 0 \qquad (i = 1, 2, 3, 4)$$

exists and we see at once that the system

$$L(y) = 0; \quad U_i(y) = 0$$

is equivalent to the integral equation

$$y(x) = \lambda \int_a^b G(x, t, \mu)\, \tau(t, \mu)\, y(t)\, d\, t,$$

or else, on setting

$$u(x) = y(x) \sqrt{\tau(x, \mu)}; \quad K(x, t, \mu) = G(x, t, \mu) \sqrt{\tau(x, \mu)\, \tau(t, \mu)},$$

to the integral equation

$$u(x) = \lambda \int_a^b K(x, t, \mu)\, u(t)\, d\, t. \tag{20}$$

It is readily seen from the known expression for the Green's function† that $K(x, t, \mu)$ is continuous in (x, t, μ) on $[(a, b), I]$. Hence the results of Weyl and Courant‡ can be applied to the equation (20). Thus the statement of Lemma 2 concerning the characteristic values is immediately proved. Let now $\lambda^{(1)}(\mu)$ be a characteristic value which tends to a characteristic value λ_0 (for $\mu = \mu_0$) as $\mu \to \mu_0$.

Let $u(x, \mu)$ be any fundamental function of (20) which corresponds to $\lambda^{(1)}(\mu)$ and for which

$$\int_a^b u^2\, d\, x = 1.$$

It has been proved by Courant that, given a positive number η arbitrarily small, a positive number $\delta_1(\eta)$ is available such that for a fixed μ which satisfies the inequality $|\mu - \mu_0| < \delta_1(\eta)$ there exists a fundamental function $u_0(x)$ of (20) (for $\lambda = \lambda_0$) such that

$$|u(x, \mu) - u_0(x)| < \eta \text{ on } (a, b); \quad \int_a^b u_0(x)^2\, d\, x = 1. \tag{21}$$

---

* These inequalities must be replaced by $\mu_1 \leqq \mu \leqq \mu_1 + h$ or by $\mu_2 - h \leqq \mu \leqq \mu_2$ if $\mu_0 = \mu_1$ or $\mu_0 = \mu_2$ respectively.

† See footnote †, p. 525.

‡ Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen, Math. Ann. 71 (1912). Courant, Zur Theorie der linearen Integralgleichungen, ibidem, 89 (1923), pp. 161–178 (174–177).

On setting

$$\varphi(x,\mu)\sqrt{\tau(x,\mu)} = u(x,\mu); \qquad \varphi_0(x)\sqrt{\tau(x,\mu_0)} = u_0(x)$$

we have

$$\varphi(x,\mu) = \lambda^{(1)}(\mu)\int_a^b G(x,t,\mu)\sqrt{\tau(t,\mu)}\,u(t,\mu)\,dt,$$

$$\varphi_0(x) = \lambda_0\int_a^b G(x,t,\mu_0)\sqrt{\tau(t,\mu_0)}\,u_0(t)\,dt.$$

Since the functions

$$\tau(x,\mu), \qquad \frac{\partial^k G(x,t,\mu)}{\partial x^k} \qquad (k = 0,\cdots,4;\ x \neq t)$$

are continuous in μ, inequalities analogous to (21) are readily established for the differences

$$\varphi^{(k)}(x,\mu) - \varphi_0^{(k)}(x) \qquad (k = 0,\cdots,4)$$

which achieves the proof of Lemma 2.

8. The following proposition furnishes a method of comparison of the characteristic values of two Sturmian systems with the same coefficients in the boundary conditions:

THEOREM 1. *In a Sturmian system let the coefficients $\varrho(x)$, $\psi(x)$, $\tau(x)$ be analytic functions of certain parameters, subject to the conditions* $1°-3°$ *of* 6. *Let the coefficients in the boundary conditions be fixed and the functions $\varrho(x)$, $\psi(x)$, $\tau(x)$ change monotonically (and no two of them simultaneously) as the parameters vary continuously. Then*

i. *When $\varrho(x)$ increases (decreases) the characteristic values do not decrease (increase).*

ii. *When $\psi(x)$ increases (decreases) the characteristic values do not decrease (increase).*

iii. *When $\tau(x)$ increases (decreases) the characteristic values which are different from zero do not increase (decrease) in absolute value.*

Let λ be a characteristic value of the Sturmian system

$$L(y) \equiv (\varrho y')' + (\psi - \lambda\tau)y = 0,$$
$$U_i(y) \equiv a_{i1}(\varrho y')' + a_{i2}\varrho y'' + a_{i3}y' + a_{i4}y]_{x=a \text{ or } b} = 0 \quad (i = 1,\cdots,4) \tag{22}$$

and $y(x)$ a corresponding normalized fundamental function. Let

$$\varrho_1(x) = \varrho(x) + \Delta\varrho, \quad \psi_1(x) = \psi(x) + \Delta\psi, \quad \tau_1(x) = \tau(x) + \Delta\tau$$

be the changed values of the functions $\varrho(x)$, $\psi(x)$, $\tau(x)$. In virtue of Lemma 2 λ will change into a characteristic value λ_1 of the Sturmian system

$$(23)\qquad \begin{aligned} \bar{L}(y) &\equiv (\varrho_1 y'')'' + (\psi_1 - \lambda_1 \tau_1) y = 0, \\ \bar{U}_i(y) &\equiv a_{i1}(\varrho_1 y'')' + a_{i2}\varrho_1 y'' + a_{i3} y' + a_{i4} y]_{x=a \text{ or } b} = 0 \quad (i = 1, \cdots, 4) \end{aligned}$$

and the increment $\Delta\lambda = \lambda_1 - \lambda$ can be made as small in absolute value as we please by taking the increments $\Delta\varrho$, $\Delta\psi$, $\Delta\tau$ sufficiently small. Moreover, under the same condition there always is found a normalized fundamental function $y_1(x)$ of the system (23), which corresponds to λ_1 and such that the absolute values of the differences

$$y_1^{(k)}(x) - y^{(k)}(x) \qquad (k = 0, \cdots, 4)$$

also are as small as we please. On combining equations (22) and (23) and integrating by parts it is readily seen that the integrated terms in the formula

$$(24)\qquad \begin{aligned} &\int_a^b \{\Delta\varrho y_1'' y'' + [\Delta\psi - (\tau\Delta\lambda + \bar{\lambda}\Delta\tau + \Delta\tau\Delta\lambda)] y_1 y\}\, dx \\ &\quad = y(\varrho_1 y_1'')' - y'\varrho_1 y_1'' + y_1'\varrho y'' - y_1(\varrho y'')']_a^b \end{aligned}$$

vanish in virtue of the boundary conditions (which, without loss of generality, may be taken in one of the canonical forms 1, i–x). Hence

$$\Delta\lambda \int_a^b \tau_1 y y_1\, dx = \int_a^b \Delta\varrho y'' y_1''\, dx + \int_a^b \Delta\psi y y_1\, dx - \bar{\lambda}\int_a^b \Delta\tau y y_1\, dx.$$

If one only of the functions $\varrho(x)$, $\psi(x)$, $\tau(x)$ is changed, according to the three cases of Theorem 1, we obtain respectively

$$(25_1)\qquad \Delta\lambda \int_a^b \tau y y_1\, dx = \int_a^b \Delta\varrho y'' y_1'\, dx;$$

$$(25_2)\qquad \Delta\lambda \int_a^b \tau y y_1\, dx = \int_a^b \Delta\psi y y_1\, dx;$$

$$(25_3)\qquad \Delta\lambda \int_a^b \tau_1 y y_1\, dx = -\bar{\lambda}\int_a^b \Delta\tau y y_1\, dx.$$

In discussing formula (25_1) we may assume for the sake of simplicity that $\varrho(x) \equiv \varrho(x, \mu)$ is an analytic function of a parameter μ such that

$$\varrho(x) = \varrho(x, \mu_0); \qquad \varrho_1(x) = \varrho(x, \mu_0 + \Delta\mu); \qquad \Delta\varrho \geqq 0 \text{ if } \Delta\mu > 0.$$

Now, given a positive ε arbitrarily small, there exists a positive number $\delta(\varepsilon)$ such that

$$|y_1^{(k)}(x) - y^{(k)}(x)| < \varepsilon \text{ if } |\Delta\mu| < \delta(\varepsilon). \qquad (k = 0, \cdots, 4).$$

Since (4,2°; 6,3°)

$$\int_a^b \tau y^2\, dx > 0,$$

we have also

$$\int_a^b \tau y y_1 \, dx > 0,$$

if ϵ is sufficiently small.

As to the right-hand member of (25_1), there are two cases to be considered, according as

$$\int_a^b \Delta\varrho \, (y'')^2 dx \equiv 0 \quad \text{or} \quad > 0 \, (\Delta\varrho \geqq 0).$$

Since the functions $\Delta\varrho$ and y'' are continuous in x, we have in the first case $y''\Delta\varrho \equiv 0$, whence

$$\int_a^b \Delta\varrho y'' y_1'' \, dx \equiv 0 \quad \text{and} \quad \Delta\lambda \equiv 0.$$

In the second case let $y'' D_\mu^s \varrho]_{\mu=\mu_0}$ be the first term of the set

$$y''\varrho, \quad y'' D_\mu \varrho, \quad y'' D_\mu^2 \varrho, \quad \cdots \qquad (\mu = \mu_0)$$

which is not identically zero on (a, b). Since $(y'')^2 \Delta\varrho \geqq 0$ when $\Delta\mu > 0$, it is readily seen that

$$(y'')^2 D_\mu^s \varrho]_{\mu=\mu_0} \geqq 0, \qquad \int_a^b D_\mu^s \varrho]_{\mu=\mu_0} (y'')^2 \, dx > 0$$

whence

$$\int_a^b \Delta\varrho \, y'' y_1'' \, dx > 0$$

provided ϵ is sufficiently small. This proves the first statement of Theorem 1; the other statements are proved in a similar fashion.

9. The application of the continuity method is based upon the use of continuous transformations of different Sturmian systems one into another. We shall be concerned exclusively with a particular case of such transformations: let $\{L_1(y), A_i^{(1)}, B_i^{(1)}\}$, $\{L_2(y), A_i^{(2)}, B_i^{(2)}\}$ be any two Sturmian systems whose boundary conditions taken in the canonical form belong to the same type among (1, i–x). Each of the given systems can be transformed continuously into another by setting up a family of Sturmian systems $\{L_\mu(y), A_i(\mu), B_i(\mu)\}$ where

$$L_\mu(y) \equiv [\varrho(x, \mu) y'']'' + [\psi(x, \mu) - \lambda \tau(x, \mu)] y$$

and

$$\varrho(x, \mu) \equiv (1-\mu')\varrho_1(x) + \mu' \varrho_2(x); \qquad \psi(x, \mu) \equiv (1-\mu'')\psi_1(x) + \mu'' \psi_2(x).$$

$$\tau(x, \mu) \equiv (1-\mu''')\tau_1(x) + \mu''' \tau_2(x);$$

$$A_i(\mu) \equiv A_i^{(1)}(1-\mu_i) + A_i^{(2)} \mu_i; \quad B_i(\mu) \equiv B_i^{(1)}(1-\mu_i) + B_i^{(2)} \mu_i.$$

As the parameters $\mu', \cdots \mu_i$ increase independently from 0 to 1 the system $\{L_1(y), A_i^{(1)}, B_i^{(1)}\}$ is transformed continuously into $\{L_2(y), A_i^{(2)}, B_i^{(2)}\}$ and vice versa. We shall call a transformation of this kind as a "μ-process".

10. We designate by $\{q, Q\}$ the class of all the Sturmian systems in which

1°. The functions $\varrho(x)$, $\tau(x)$ are positive and bounded, $\psi(x)$ is bounded in absolute value and

$$0 < q \leqq \varrho(x) \leqq Q,$$
$$0 < q \leqq \tau(x) \leqq Q,$$
$$|\psi(x)| \leqq Q,$$

where q and Q are positive constants.

2°. In the canonical form (1, i–x) of the boundary conditions the coefficients A_i, B_i do not exceed Q in absolute value.

The Sturmian systems of any class $\{q, Q\}$ have an important property in common:

THEOREM 2. *The characteristic values of all the Sturmian systems of the class $\{q, Q\}$ are uniformly bounded below, in the sense that there exists a positive constant $A \equiv A(q, Q)$ which depends only on q and Q and such that the characteristic values of any system of the class $\{q, Q\}$ are $> A$.*

Let $\{q, Q\}_0$ be the subclass of $\{q, Q\}$, in which the operator $L(y)$ has the particular form

$$L_0 y \equiv (qy')' + (-Q - \lambda q)y.$$

The characteristic values of all the systems of $\{q, Q\}_0$ are uniformly bounded below. On the other hand, if $\{L(y), A_i, B_i\}$ be any given system of $\{q, Q\}$ we can always find a system of $\{q, Q\}_0$ with the same coefficients A_i, B_i in the boundary conditions. On applying the „μ-process" to pass from $\{L(y), A_i, B_i\}$ to $\{L_0(y), A_i, B_i\}$ we see that the functions $\varrho(x, \mu)$, $\psi(x, \mu)$ increase while $\tau(x, \mu)$ decreases. Since the negative characteristic values can not decrease (Theorem 1) they remain uniformly bounded below.

COROLLARY. *No Sturmian system* has $-\infty$ as a limiting point of the characteristic values; the number of the negative characteristic values of any such system is finite.*

If $\{L(y), A_i, B_i\}$ be any Sturmian system in question, let

$$0 < q \leqq \varrho(x) \leqq Q; \quad 0 \leqq \tau(x) \leqq Q; \quad |\psi(x)| \leqq Q.$$

On setting

$$L_1(y) \equiv (\varrho y')' + (\psi - \lambda Q)y$$

and applying the „μ-process" to pass from $\{L_1(y), A_i, B_i\}$ to $\{L(y), A_i, B_i\}$ the characteristic values will move away from the origin (Theorem 1),

* Cfr. the beginning of 2.

whence the number of the negative characteristic values can not increase **and remains finite. This number will actually decrease when some of the characteristic values are lost at $\lambda = -\infty$.**

11. THEOREM 3. *Let* $\{L_1(y), A_i^{(1)}, B_i^{(1)}\}$ *and* $\{L_2(y), A_i^{(2)}, B_i^{(2)}\}$ *be any two regular Sturmian systems in which the coefficients* $\tau_1(x)$ *and* $\tau_2(x)$ *are* >0, *and the canonical forms of the boundary conditions are of the same type* (1, i–x). *If one of the given systems be transformed into another by the „μ-process", then the characteristic values remain finite and none of them is lost or gained at* $\lambda = \pm\infty$ (in virtue of Lemma 2 all the characteristic values will vary continuously with the μ's).

It has been already proved (Theorem 2) that the characteristic values remain uniformly bounded below. It remains now to prove that none of them can $\to +\infty$ as the μ's vary. We first observe that, at any stage of the „μ-process", the system $\{L_\mu(y), A_i(\mu), B_i(\mu)\}$ has infinitely many real characteristic values (4, 3° — 4°); all of them, except perhaps a finite number, satisfy the condition of Lemma 1 and hence they are simple (5, Corollary to Lemma 1) and can not coincide during the „μ-process".

If, now, some of the characteristic values $\to +\infty$ as the μ's approach certain limiting values μ_0's in (0, 1), let $\lambda(\mu)$ be the algebraically smallest among those λ's. Since none of the λ's which are $>\lambda(\mu)$ can coincide with $\lambda(\mu)$ as $(\mu)\to(\mu_0)$, all these λ's also will $\to +\infty$, with the result that our Sturmian system has only a finite number of characteristic values at $(\mu) = (\mu_0)$, which is not possible. Hence Theorem 3 is proved.

12. The discussion of the oscillation properties of the regular Sturmian systems is based upon the following.

LEMMA 3. *Let any regular Sturmian system be transformed into another regular system by the "μ-process" under the conditions of Theorem* 3. *The fundamental functions corresponding to the characteristic values which for all values of the μ's in* (0, 1) *satisfy the condition of Lemma* 1 (*or that of the Remark to Lemma* 1 *at the end of* 5, *in the case where* $\psi(x, \mu) \equiv 0$) *preserve their oscillation properties. In other words, the number of zeros of any such fundamental function interior to* (a, b) *remains the same for all values of the μ's in* (0, 1).

Let $\varphi(x, \mu)$ be any fundamental function in question. Since the corresponding characteristic value $\lambda(\mu)$ remains simple (Corollary to Lemma 1) $\varphi(x, \mu)$ is continuous in the μ's (Remark to Lemma 2). The function $\varphi(x, \mu)$ has a finite number of zeros interior to (a, b) (4, Corollary to 2°), which, being simple (Lemma 1) also are continuous. The function $\varphi(x, \mu)$ may have a root at one (or at both) of the end-points a, b (the case where $\varphi(a, \mu) \neq 0$ or $\varphi(b, \mu) \neq 0$ being considered as a root of multiplicity zero). However, by virtue of Lemma 1, the multiplicity of such a root remains

invariant under our "μ-process". Hence the number of zeros of $\varphi(x, \mu)$ interior to (a, b) is also invariant and Lemma 3 is proved.

13. *Remark.* The definition of the "μ-process" can be immediately extended to the case where the given Sturmian system is completely regular and the boundary conditions are taken in the forms (xi), with the agreement of the footnote *, p. 529, concerning the case where $C_1 = \infty$ or $D_1 = \infty$.

Theorem 3 *and Lemma* 3 *hold true for this modified form of the "μ-process"* (*and the completely regular systems*).

The proof of Theorem 3 is extended without any changes. Some modifications are necessary, however, in the proof of Lemma 3, since the multiplicity of a root of $\varphi(x, \mu)$ at $x = a$ $(x = b)$ is different in the two cases: $C_0 + C_2 \neq 0$ $(D_0 + D_2 \neq 0)$ and $C_0 = C_2 = 0$ $(D_0 = D_2 = 0)$.

The essential features of the proof are sufficiently illustrated by the discussion of one particular case: the two given completely regular Sturmian systems have the same operators

$$L(y) \equiv (\varrho y'')'' + (\psi - \lambda\tau)y = 0$$

subject to the condition of Lemma 1 (or to that of the Remark to Lemma 1) the boundary conditions for one of them being

$$A_6^{(1)}(y) \equiv \varrho y'' - C_1 y']_{x=a} = 0; \qquad A_7^{(1)}(y) \equiv y(a) = 0, \qquad 0 \leqq C_1 < \infty$$
$$B_6^{(1)}(y) \equiv B_6(y) = 0; \qquad B_7^{(1)}(y) \equiv B_7(y) = 0$$

and for the other

$$A_6^{(2)}(y) \equiv \varrho y'' - C_0(\varrho y'')' - C_1 y']_{x=a} = 0;$$
$$A_7^{(2)}(y) \equiv y + C_2(\varrho y'')' - C_0 y']_{x=a} = 0;$$
$$B_6^{(2)}(y) \equiv B_6(y) = 0; \quad B_7^{(2)}(y) \equiv B_7(y) = 0; \quad C_0 + C_2 \neq 0.$$

In applying the "μ-process" to pass from the first system to the second we set up the system

$$L(y) = 0,$$
$$\varrho y'' - \mu C_0(\varrho y'')' - C_1 y']_{x=a} = 0; \quad y + \mu C_2(\varrho y'')' - \mu C_0 y']_{x=a} = 0, \tag{26}$$
$$B_6(y) = 0; \qquad B_7(y) = 0.$$

Let $\varphi(x, \mu)$ be the fundamental function of (26) which corresponds to a characteristic value $\lambda(\mu)$. The condition of Lemma 1 (Remark to Lemma 1) being satisfied, $\varphi(x, \mu)$ and its roots are continuous in μ. By virtue of Lemma 1 we have here

$$\varphi(a, 0) = 0; \qquad \varphi'_x(a, 0) \neq 0,$$
$$\varphi(a, \mu) \neq 0 \quad \text{for} \quad 0 < \mu \leqq 1.$$

Without loss of generality we may assume $\varphi'_x(a, 0) > 0$. Then necessarily $(\varrho\varphi''_x)'_{x=a,\mu=0} < 0$, since a contrary hypothesis would imply a contradiction to the boundary conditions at $x = b$.* Thus

$$\varphi'_x(a, \mu) > 0, \quad (\varrho\varphi''_x)'\big]_{x=a} < 0 \quad \text{for} \quad 0 \leqq \mu \leqq h \tag{27}$$

where h is a fixed positive number sufficiently small. It follows then from (27) that $\varphi(x, \mu) > \varphi(a, \mu) > 0$, $0 < \mu \leqq h$, $a \leqq x \leqq a + \varepsilon$ where ε is a fixed positive number which depends only on h. Hence, as μ increases from zero, a root at $x = a$ is lost by the function $\varphi(x, \mu)$, but no roots are gained in the neighborhood of $x = a$. Hence the number of roots interior to (a, b) is invariant under our "μ-process", which is the desired result. Other cases are treated in an analogous fashion.

14. Now we are in a position to investigate the oscillation properties of the fundamental functions of regular Sturmian systems.

We begin with a completely regular system $(\mathfrak{C})$ for which $\psi(x) \equiv 0$ and $\tau(x) > 0$:

$$\begin{aligned} L(y) &\equiv (\varrho y'')'' - \lambda\tau y = 0, \\ A_6(y) &= 0, \quad A_7(y) = 0, \quad C_1 \neq \infty, \quad D_1 \neq \infty, \\ B_6(y) &= 0, \quad B_7(y) = 0. \end{aligned}$$

All the characteristic values of $(\mathfrak{C})$ are $\geqq 0$ (4, Corollary to 5°). We can prove, however, that they are > 0. To do this we set

$$\varrho_0 = \min \varrho(x); \qquad \tau_0 = \min \tau(x); \qquad T = \max \tau(x) \text{ on } (a, b)$$

and introduce an auxiliary system $(\mathfrak{C}_0)$:

$$\begin{aligned} &\varrho_0 y^{\mathrm{IV}} - \lambda T y = 0, \\ &\varrho_0 y'' - C_0\varrho_0 y''' - C_1 y']_{x=a} = 0; \quad y + C_2\varrho_0 y''' - C_0 y']_{x=a} = 0, \\ &\varrho_0 y'' + D_0\varrho_0 y''' + D_1 y']_{x=b} = 0; \quad y - D_2\varrho_0 y''' + D_0 y']_{x=b} = 0 \end{aligned} \tag{28}$$

from which the system $\mathfrak{C}$ can be a suitable "μ-process".

System $(\mathfrak{C}_0)$ being completely regular has all the characteristic values $\geqq 0$. On the other hand $\lambda = 0$ can not be a characteristic values of $(\mathfrak{C}_0)$, since, if it were, the corresponding fundamental function must be of form

$$y = \alpha_0 x^3 + \alpha_1 x^2 + \alpha_2 x + \alpha_3.$$

On substituting in the boundary conditions (28) we obtain a system of 4 linear homogeneous equations in $\alpha_0, \cdots, \alpha_3$, whose determinant is > 0 provided the conditions $C_0, C_1, C_2, D_0, D_1, D_2 \geqq 0$ are fulfilled.†

* Cfr. 5, proof of Lemma 1.

† The computation is simplified if we take, without loss of generality, $a = 0$, $b = 1$; $\varrho_0 = 1$, and replace α_0 by $\alpha_0/6$ and α_1 by $\alpha_1/2$.

If now, we apply the "μ-process" to pass from $(\mathfrak{C}_0)$ to $(\mathfrak{C})$ the characteristic values can not decrease (Theorem 1) which proves that $\lambda = 0$ is not a characteristic value of the system $(\mathfrak{C})$. Moreover, the Remark of 13 can be applied in the present case with the result that the oscillation properties of the fundamental functions do not change in the transition from $(\mathfrak{C}_0)$ to $(\mathfrak{C})$ by the "μ-process". System $(\mathfrak{C})$ is also obtained by the "μ-process" from the simpler system $(\mathfrak{C}')$

$$\varrho_0 y^{IV} - \lambda T y = 0,$$
$$y(a) = 0; \quad y''(a) = 0; \quad y(b) = 0; \quad y''(b) = 0.$$

It is known* that $(\mathfrak{C}')$ is possessed of infinitely many characteristic values (simple and positive)

$$0 < \lambda_0^{(1)} < \lambda_1^{(1)} < \cdots < \lambda_\nu^{(1)} < \cdots$$

and also that the fundamental function $\varphi_\nu^{(1)}(x)$ which corresponds to $\lambda_\nu^{(1)}$ has exactly ν simple roots in the interior of (a, b). In virtue of Remark of 13, these properties are preserved in the transition from $(\mathfrak{C}')$ to $(\mathfrak{C})$ by the "μ-processes".

The case where one or both of the coefficients C_1, D_1 become ∞ may be treated in entirely analogous fashion. Thus the following oscillation theorem is established:

THEOREM 4. *A completely regular Sturmian system in which* $\psi(x) \equiv 0$ *and* $\tau(x) > 0$ *has infinitely many simple and positive characteristic values*

$$0 < \lambda_0 < \lambda_1 < \cdots < \lambda_\nu < \cdots.$$

The fundamental function $\varphi_\nu(x)$ *which corresponds to* λ_ν *has exactly* ν *simple roots in the interior of* (a, b).

15. A regular Sturmian system $(\mathfrak{R})$ in which $\psi(x) \equiv 0$ and $\tau(x) > 0$ can be obtained by the "μ-process" from a completely regular system $(\mathfrak{C})$ with the same operator $L(y)$. In the transition from $(\mathfrak{C})$ to $(\mathfrak{R})$, only a finite number among the characteristic values may pass through $\lambda = 0$ and become negative (Theorem 3). We designate them as "irregular" and the remaining characteristic values (which never approach zero) as "regular" ones. Since Lemm 3 is applicable to the regular characteristic values, the oscillation properties of the corresponding "regular" fundamental functions will not change.

Now, it is readily proved that the number of the irregular characteristic values can not exceed 4, so that $\lambda_4, \lambda_5, \cdots$ are surely regular (and as

* Haupt, loc. cit., pp. 39–50.

such are simple and distinct from the irregular characteristic values. Indeed, let $(\mathfrak{R}_0)$ be the regular system which is obtained from $(\mathfrak{R})$ by replacing $\varrho(x)$ by $\varrho_0 = \min \varrho(x)$ and $\tau(x)$ by $T = \max \tau(x)$. Let $(\mathfrak{C}_0)$ be the completely regular system which is obtained in the same fashion from $(\mathfrak{E})$.

In the transition from $(\mathfrak{C}_0)$ to $(\mathfrak{R}_0)$ by the "μ-process" all the characteristic values $\lambda_k^{(1)}(\mu)$ are regular at the beginning, hence the fundamental function $\varphi_4^{(0)}(x, \mu)$ has four simple roots in the interior of (a, b). Now, if at a certain stage of the "μ-process" 0 becomes a characteristic value, the corresponding fundamental function (or functions in case where 0 is a multiple characteristic value) is a polynomial of the degree $\leqq 3$. This shows that $\lambda_4^{(0)}(\mu)$ can not $\to 0$, for otherwise the function $\varphi_4^{(0)}(x, \mu)$ which has four roots in (a, b) would have as a limiting function a polynomial of the degree $\leqq 3$, which is not possible.

Now we pass from $(\mathfrak{R}_0)$ to $(\mathfrak{R})$ by the "μ-process". Since the coefficient $\tau(x, \mu) \equiv (1-\mu)\, T + \mu\, \tau(x)$ decreases, the positive characteristic values can not decrease (Theorem 1). Hence $\lambda_4(\mu), \cdots$ remain positive, finite and simple and the corresponding fundamental functions do not change their oscillations properties (Theorem 3 and Lemma 3).

This result may be extended to any regular Sturmian system $\mathfrak{S}$ with $\psi(x) \equiv 0$ (and $\tau(x) \geqq 0$ and $\not\equiv 0$ on any subinterval of (a, b), according to what has been said in 2). Let $(\mathfrak{S}_0)$ be the regular system which is obtained from $(\mathfrak{S})$ by replacing $\tau(x)$ by $T = \max \tau(x)$. If we pass from $(\mathfrak{S}_0)$ to $(\mathfrak{S})$ by the "μ-process", the coefficient $\tau(x, \mu)$ decreasing, the positive characteristic values can not decrease (Theorem 1), and the negative characteristic values can not increase whence their number does not exceed 4. Then the same argument as the one used in the proof of Theorem 3 will show that no characteristic value can $\to +\infty$. Hence the characteristic values $\lambda_4(\mu), \cdots$ remain positive, finite and simple and (Lemma 3) the corresponding fundamental functions do not change their oscillation properties. On collecting all these results we obtained the

THEOREM 5. *A regular Sturmian system in which $\psi(x) \equiv 0$ has infinitely many characteristic values which are real and, except for at most four $(\lambda_{3-\sigma},\ 0 \leqq \sigma \leqq 3)$, positive and simple. On denoting them by*

$$0 < \lambda_4 < \lambda_5 < \cdots < \lambda_\nu < \cdots$$

the fundamental function $\varphi_\nu(x)$ which corresponds to λ_ν has exactly ν simple roots in the interior of (a, b).

16. The discussion of the most general regular Sturmian system (with $\psi(x) \not\equiv 0$) will be carried out under the assumption $\tau(x) > 0$.

Let $(\mathfrak{T})$ be the given regular system. On setting

$$\begin{aligned} \psi_0 &= \min \psi(x); & \Psi &= \max \psi(x) \\ \tau_0 &= \min \tau(x); & T &= \max \tau(x) \end{aligned} \quad \text{on} \quad (a, b)$$

denote by $(\mathfrak{T}_0)$ the regular system obtained from $(\mathfrak{T})$ by replacing $\psi(x)$ by ψ_0 and $\tau(x)$ by T. The substitution

$$\lambda' = \lambda T - \psi_0$$

transforms $(\mathfrak{T}_0)$ into an equivalent regular system $(\mathfrak{T}')$ of the type to which Theorem 5 may be applied. Let

$$\lambda'_{3-\sigma} (0 \leqq \sigma \leqq 3), \qquad \lambda'_4, \lambda'_5, \cdots$$

be the characteristic values of $(\mathfrak{T}')$ and

$$\lambda_k^{(0)} = \frac{\lambda'_k + \psi_0}{T}$$

the corresponding characteristic values of $(\mathfrak{T}_0)$. Of these, $\lambda_4^{(0)}, \cdots$ surely are possessed of the oscillation properties as indicated by Theorem 5.

We pass now from $(\mathfrak{T}_0)$ to $(\mathfrak{T})$ by the "μ-process". Since the coefficient $\psi(x, \mu) \equiv (1-\mu'')\psi_0 + \mu''\psi(x)$ increases and $\tau(x, \mu) \equiv (1-\mu''')T + \mu'''\tau(x)$ decreases, the positive characteristic values can not decrease (Theorem 1). The condition of Lemma 1

$$\psi(x, \mu) - \lambda(\mu)\tau(x, \mu) < 0$$

is a fortiori satisfied by the characteristic values

$$\lambda_m(\mu), \qquad \lambda_{m+1}(\mu), \cdots$$

where m is the greatest of the two numbers m_0 and 4, and m_0 is defined by:

$$\lambda'_{m_0} > \frac{\Psi T - \psi_0 \tau_0}{\tau_0} \geqq \lambda'_{m_0-1}, \quad \lambda_{m_0}^{(0)} > 0.$$

We therefore obtain the following oscillation theorem:

THEOREM 6. *A regular Sturmian system in which $\tau(x) > 0$ has infinitely many characteristic values which all are real. With exception of no more than m of them, where the integer m is defined above, the characteristic values are simple and positive. If we denote them by*

$$0 < \lambda_m < \lambda_{m+1} < \cdots,$$

then the fundamental function $\varphi_\nu(x)$ which corresponds to λ_ν has exactly ν simply roots in the interior of (a, b).

17. A curious example in which the "irregular" characteristic values occur is furnished by the regular Sturmian system (2[)

$$L(y) \equiv y^{IV} - \lambda y = 0,$$
$$y'''(0) + A_1 y(0) = 0; \quad y''(0) = 0,$$
$$y'''(1) = 0; \quad y''(1) = 0.$$

For this system $\lambda = 0$ is always a characteristic value, simple if $A_1 \neq 0$ and double if $A_1 = 0$ with the (non-normalized) fundamental functions x or x, 1, respectively.

The „Dirichlet's identity"

$$\int_0^1 y\, L(y)\, dx = \int_0^1 (y'')^2\, dx - \lambda \int_0^1 y^2\, dx + A_1 y(0)^2$$

shows at once that in the case where $A_1 > 0$ all the characteristic values $\neq 0$ are positive.

In the case $A_1 = 0$ it has been shown by Haupt* that there exist infinitely many characteristic values

$$0 = \lambda_0^{(0)} = \lambda_1^{(0)} < \lambda_2^{(0)} < \cdots < \lambda_\nu^{(0)} < \cdots$$

and that the fundamental function corresponding to $\lambda_\nu^{(0)}$ has ν simple roots in the interior of (0, 1). Then, by using Lemma 3, it is readily seen that if $A_1 > 0$, our system is possessed of infinitely many real characteristic values

$$0 = \lambda_0 < \lambda_1 < \cdots < \lambda_\nu < \cdots$$

which are simple and such that the fundamental function $\varphi_\nu(x)$ which corresponds to λ_ν, has ν simple roots in the interior of (0, 1).

The same conclusion is true in the case $A_1 < 0$ of the characteristic values $\lambda_2, \lambda_3, \cdots$ since in passing from the case $A_1 = 0$ to the case $A_1 < 0$ by the „μ-process", $\lambda_2(\mu)$ can not $\to 0$ (the corresponding fundamental function $\varphi_2(x, \mu)$ has two roots in (0, 1) and can not have a linear function as a limiting function). On the other hand $\lambda = 0$ is a characteristic value of (2[). Hence (2[) can have no more than one negative characteristic value. That it actually has one, may be shown as follows: On setting $\lambda = -4\sigma^4$ we shall prove that the system

$$y^{IV} + 4\sigma^4 y = 0,$$
$$y'''(0) + A_1 y(0) = 0; \quad y''(0) = 0; \quad y'''(1) = 0; \quad y''(1) = 0$$

has at least one positive characteristic value if $A_1 < 0$.

* See footnote ‡, p. 521.

A fundamental system of solutions of $y^{IV}+4\sigma^4 y = 0$ is:

$$y_1(x) = \operatorname{ch}\sigma x \cos\sigma x; \quad y_2(x) = \operatorname{ch}\sigma x \sin\sigma x; \quad y_3(x) = \operatorname{sh}\sigma x \cos\sigma x; \quad y_4(x) = \operatorname{sh}\sigma x \sin\sigma x$$

and an easy computation shows that the characteristic values $\neq 0$ are determined from the equation

$$\Delta(\sigma) \equiv 2\sigma^3(\operatorname{sh}^2\sigma - \sin^2\sigma) + A_1(\operatorname{sh}\sigma \operatorname{ch}\sigma - \sin\sigma\cos\sigma) = 0.$$

The existence of at least one positive root of this equation is assured by the fact that $\Delta(\sigma)>0$ for large positive values of σ and $\Delta(\lambda) = \frac{4}{3}A_1\sigma^3 + \cdots < 0$ for small positive σ. Thus (2l) is possessed of the characteristic values

$$\lambda_0 < 0; \quad \lambda_1 = 0, \lambda_2, \lambda_3 \ldots.$$

If $\varphi_\nu(x)$ denotes the fundamental function which corresponds to λ_ν, then $\varphi_0(x)$, being orthogonal to $\varphi_1(x) = x$, must have at least one root, $\varphi_1(x)$ has no roots and $\varphi_\nu(x)$ $(\nu \geqq 2)$ has exactly ν simple roots in the interior of (0, 1). This does not agree with a proposition stated by Courant.*

* „Ordnet man die Eigenfunktionen einer sich selbst adjungierten Differentialgleichung $L(u) + \lambda\varrho u = 0$, $\varrho > 0$, für ein Gebiet G bei beliebigen homogenen Randbedingungen nach wachsenden Eigenwerten, so teilt die n-te Eigenfunktion u_n durch ihre Nullstellen das Gebiet in nicht mehr als n Teilgebiete. Dabei werden über die Ordnung der Differentialgleichung sowie über die Anzahl der unabhängigen Veränderlichen keinerlei Voraussetzungen gemacht", Courant-Hilbert, Methoden der Mathematischen Physik, vol. 1, Berlin, Springer, 1924, p. 365. Also: „Die erste Eigenfunktion eines Eigenwertproblems keine Knoten im Innern des Grundgebietes besitzen kann", ibidem, p. 364. Cfr. also: Ein allgemeiner Satz zur Theorie der Eigenfunktionen selbstadjungierter Differentialausdrücke, Nachr. Ges. der Wiss. Gött., Math.-phys. Klasse, 1923, pp. 81–84.

LENINGRAD,
January, 1927.

A NOTE ON STIELTJES INTEGRALS.*

By Ray N. Haskell.

1. In a recent paper by Professors H. E. Bray and G. C. Evans† the existence of a Stieltjes integral with reference to a bounded, additive function of regular curves w was discussed in the case that the discontinuities of that function were regular. But the question remains open as to whether the integral exists, as a Riemann sum, if the function $F(w)$ with respect to which the integration is carried out is unrestricted as to its discontinuities. That question is to be answered affirmatively in the following paper.

We take as our fundamental region Δ, the recangle $(0 \leqq x \leqq 1, 0 \leqq y \leqq 1)$ and consider first a mode of subdivision of Δ into the polygons π of a polygonal lattice. A finite polygonal pluricell π consists of a finite number of non-overlapping polygons π_i in Δ,

$$\pi = \pi_1 + \pi_2 + \cdots + \pi_n$$

with or without common boundary points. If for all such pluricells the conditions

$$|F(\pi)| < K,$$
$$F(\pi) = F(\pi_1) + F(\pi_2) + \cdots + F(\pi_n),$$

are satisfied, $F(\pi)$ is said to be a *bounded additive function of polygonal pluricells* in Δ.

Let $P(\pi)$ be the upper bound of $F(\pi')$ for all π' contained‡ in π, and similarly $-N(\pi)$ the lower bound of $F(\pi')$. We remark first that $P(\pi)$ and $N(\pi)$ are both bounded and additive functions of finite pluricells (whence it follows as in the cited memoirs that they are *completely* additive), and that

$$F(\pi) = P(\pi) - N(\pi).$$

---

* Received January 24, 1928. Presented to the American Mathematical Society, May 1927. The discussion of this question was suggested by Professor H. E. Bray.

† H. E. Bray and G. C. Evans, "A Class of Functions Harmonic within the Sphere", Amer. Journ. Math., vol. 49 (1927), pp. 153–180.

‡ The analogous questions are discussed for *plurisegments* in two dimensions by A. J. Maria, "Functions of Plurisegments", Trans. Amer. Math. Soc., vol. 28 (1926), pp. 448–471. This paper is a generalization of Vitali's memoir, Analisi delle funzioni a variazione limitata, Rend. Circ. Mat. Palermo, vol. 46 (1922), p. 388. The reader will find in Maria's article the definitions of "contained in" etc. for finite and infinite plurisegments.

It is obvious that $P(\pi)$ is bounded. In order to prove $P(\pi)$ is additive, suppose for simplicity that

$$\pi = \pi_1 + \pi_2.$$

Evidently

$$P(\pi) \geqq P(\pi_1) + P(\pi_2) \tag{1}$$

since any finite pluricell contained in π_1 added to any finite pluricell contained in π_2 is contained in π.

Now we can find a finite pluricell π_{12} in π for which

$$F(\pi_{12}) > P(\pi) - \varepsilon$$

for any $\varepsilon > 0$ given in advance. This pluricell will be divided into two finite pluricells π_1' and π_2' contained in π_1 and π_2 such that

$$F(\pi_1') + F(\pi_2') > P(\pi) - \varepsilon$$

therefore

$$P(\pi_1) + P(\pi_2) \geqq P(\pi). \tag{2}$$

The statement that π_{12} will be divided into two finite π_1 and π_2 depends on the fact that polygonal lines intersect each other only a finite number of times. The same is not true for regular curves in general, and therefore this method of proof is not directly applicable to them. From (1) and (2)

$$P(\pi) = P(\pi_1) + P(\pi_2)$$

and in similar fashion if $\pi = \sum_1^n \pi_i$

$$P(\pi) = \sum_1^n P(\pi_i).$$

We notice further that if

$$R(\pi) = P(\pi) - F(\pi)$$

$-R(\pi)$ will be the lower bound of $F(\pi')$ for all π' contained in π. Consequently $R(\pi) = N(\pi)$ and

$$F(\pi) = P(\pi) - N(\pi).$$

2. Now given $h(Q)$ a continuous function of the point Q, and therefore uniformly continuous in Δ, form the two sums

$$S = \sum_i M(\pi_i)\, P(\pi_i),$$

$$s = \sum_i m(\pi_i)\, P(\pi_i)$$

where $M(\pi_i)$ and $m(\pi_i)$ are respectively the greatest and least values of $h(Q)$ in π_i. Evidently

$$m(\Delta)\,P(\Delta) \leqq s \leqq S \leqq M(\Delta)\,P(\Delta). \tag{3}$$

If we make a further subdivision $\pi_i = \sum \pi_{ij}$ and let s' and S' be the corresponding Riemann sums,

$$S' = \sum_{ij} M(\pi_{ij})\,P(\pi_{ij}),$$
$$s' = \sum_{ij} m(\pi_{ij})\,P(\pi_{ij}),$$

we shall have

$$s' \geqq s, \quad S' \leqq S.$$

If we make any other polygonal lattice on Δ, and let s and S' be again the corresponding Riemann sums, we shall have

$$S' \geqq s, \quad s' \leqq S.$$

For if, for example, we suppose $S' < s$ and superimpose the two modes of subdivision, we shall have a polygonal lattice which will be a subdivision of the lattice which yields the sums s, S and the lattice which yields the sums s', S'. If the corresponding sums for the new lattice are denoted by s'', S'' we should have

$$S'' \leqq S' < s \leqq s''.$$

But by (3),

$$S'' \geqq s''.$$

The sums S have a lower bound $\bar{I}$ and the sums s have an upper bound $\underline{I}$. Evidently

$$\bar{I} \geqq \underline{I},$$

for otherwise we should have some s greater than some S. Now

$$S - s = \sum_i [M(\pi_i) - m(\pi_i)]\,P(\pi_i) \leqq [\max (M(\pi_i) - m(\pi_i))]\,P(\Delta).$$

Let δ be the diameter of the largest π_i in Δ. Given $\epsilon > 0$, there is a δ small enough so that

$$\max [M(\pi_i) - m(\pi_i)] < \frac{\epsilon}{P(\Delta)}$$

because of the uniform continuity of $h(Q)$ in Δ. Therefore

$$0 \leqq \bar{I} - \underline{I} \leqq S - s < \varepsilon,$$

and

$$\bar{I} = \underline{I} = I.$$

This value is the Stieltjes integral of $h(Q)$ with respect to $P(\pi)$ and is written as

$$I = \int_{\Delta} h(Q)\, dP(\pi_Q).$$

That this integral is the limit of a Cauchy sum follows from the inequalities,

$$s \leqq I \leqq S,$$
$$s \leqq \sum_i h(Q_i)\, P(\pi_i) \leqq S,$$

whence

$$I = \lim_{\delta=0} \sum_i h(Q_i)\, P(\pi_i).$$

Moreover we have

$$\lim_{\delta=0} \sum_i h(Q_i)\, F(\pi_i) = \lim_{\delta=0} \sum_i h(Q_i)\, P(\pi_i) - \lim_{\delta=0} \sum_i h(Q_i)\, N(\pi_i),$$

and therefore, if we denote the left hand member by the symbol $\int_{\Delta} h(Q)\, dF(\pi_Q)$, we have

$$\int_{\Delta} h(Q)\, dF(\pi_Q) = \int_{\Delta} h(Q)\, dP(\pi_Q) - \int_{\Delta} h(Q)\, dN(\pi_Q).$$

3. We wish now to consider the Cauchy sum with reference to a lattice L_w of regular curves.* As in the figure, Δ is devided into a finite number of simply connected regions by regular curves w. Using Prof. Veblen's† terminology we call a point of intersection of two or more curves a 0-cell, the curves themselves 1-cells and the regions 2-cells. From an interior point of each 2-cell, draw a polygonal line to a regular point (not a 0-cell), of each 1-cell, and for two contiguous 2-cells choose the same regular point. The resultant lattice we call $L_{w|\pi}$. The polygonal lines thus formed will be finite in number and if the curves w are deleted we shall have a polygonal lattice L_π as in §§ 1 and 2. We could equally well start with the polygons as given and divide them with the same regular curves and get the same configuration as the above. We denote this resultant lattice by $L_{\pi|w}$. We have, then,

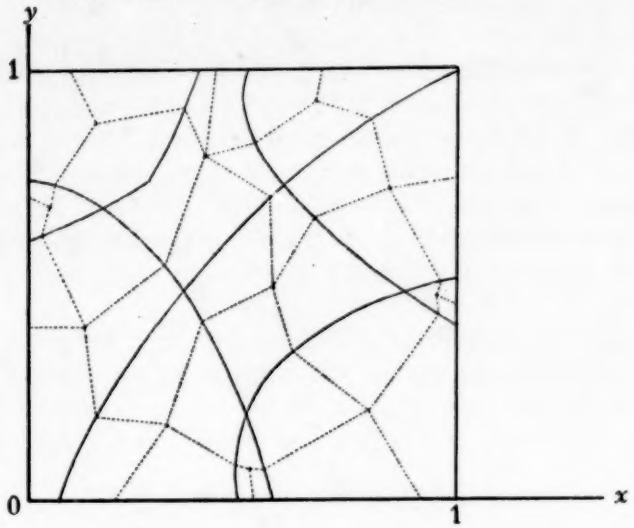
$$L_{\pi|w} = L_{w|\pi}.$$

* See Bray and Evans, loc. cit., p. 169.

† Oswald Veblen, Analysis Situs, Cambridge Colloquium Lectures, part II, p. 41.

For simplicity in notation we denote by $(\pi|w)$ a cell obtained by subdividing a polygonal cell π by regular arcs w of the lattice L_w, and by $(w|\pi)$ a cell obtained by subdividing a regular cell w by polygonal lines of the lattice L_π. By $Q(w)$ we mean any point of the cell w. We have, then,

$$h(Q(\pi|w)) = h(Q(\pi)) + \eta_1(\pi|w),$$
$$h(Q(w|\pi)) = h(Q(w)) + \eta_2(w|\pi),$$



where $|\eta_1(\pi|w)|$, $|\eta_2(w|\pi)|$ may be made less than $\varepsilon/2K$, K given, for a sufficiently small value of δ, if δ is the maximum diameter of the cells w of L_w. Now

$$\begin{aligned}
&\left|\sum h(Q(w))\,F(w) - \sum\sum h(Q(w|\pi))\,F(w|\pi)\right| \\
&= \left|\sum h(Q(w))\,F(w) - \sum\sum [h(Q(w)) + \eta_1(w|\pi)]\,F(w|\pi)\right| \\
&= \left|\sum h(Q(w))\,F(w) - \quad \sum h(Q(w))\,F(w) - \sum\sum \eta_1(w|\pi)\,F(w|\pi)\right| \\
&\leqq \frac{\varepsilon}{2K}\sum |F(w)| < \frac{\varepsilon}{2},
\end{aligned} \tag{4}$$

for, since $F(w)$ is a bounded additive function of regular curves, there will be a constant K such that $\sum|F(w)| < K$ for every finite set of non-overlapping regular curves.

In the same way,

$$\left|\sum h(Q(\pi))\,F(\pi) - \sum\sum h(Q(\pi|w))\,F(\pi|w) < \frac{\varepsilon}{2}. \tag{5}$$

But each $(\pi|w)$ can be identified with some $(w|\pi)$ and therefore from (4) and (5) if we take $Q(\pi|w)$ as the point $Q(w|\pi)$, we have

$$\left|\sum h(Q(w))\,F(w) - \sum h(Q(\pi))\,F(\pi)\right| < \varepsilon,$$

and therefore

$$\lim_{\delta=0} \sum h(Q(w))\,F(w)$$

exists and is equal to

$$\int_\Delta h(Q)\,dF(\pi_Q).$$

We have therefore proved that

$$\int_\Delta h(Q)\,dF(w_Q) = \lim_{\delta=0} \sum h(Q(w))\,F(w) = \int_\Delta h(Q)\,dF(\pi_Q),$$

provided that $h(Q)$ is continuous in Δ, and $F(w)$ is a bounded additive function of regular curves w.

THE RICE INSTITUTE,
HOUSTON, TEXAS.

ON MOTIONS WHICH SATISFY KEPLER'S FIRST AND SECOND LAWS.*

By L. R. Ford.

1. **Introduction.** If a particle P move in a plane in such a manner that a line segment joining the particle to a fixed point O of the plane sweeps out equal areas in equal times, that is, so that Kepler's first law is satisfied, it is well known that the resultant acceleration on the particle is directed constantly along the line OP; in other words, the force is central. Conversely motion under a central force obeys the law of areas. By Kepler's second law we understand the requirement that the path of P be a conic with O as a focus. If the second law is satisfied for all initial conditions and if the acceleration toward O is a function of OP alone, it has been proved that the acceleration is inversely proportional to the square of the distance.

In the present paper we shall investigate the case in which the central acceleration is a function of the distance and of the velocity. We shall prove that the most general central acceleration, φ, such that all motions satisfy the second law is given by the formula

$$F(\varphi r^2,\ 2r\varphi - v^2) = 0,$$

where r is the distance OP, v is the velocity of P, and F is an arbitrary function. The character of the orbit is investigated, and the relation between the period and the mean distance in the case of elliptical orbits is determined. It is shown that Kepler's third law may be replaced by any functional relation between the period and the mean distance and a suitable φ is thereby determined.

In the treatment of the problem use is made of the determinant known as the Wronskian. The Wronskian, which seems not to have been employed hitherto in problems of this sort, furnishes a very direct proof that the orbit of a planet under the Newtonian law is a conic — the so-called Fundamental Theorem of Celestial Mechanics.

The dependence of the acceleration upon the velocity as well as the distance is suggested by modern physical theories wherein various physical entities are dependent upon the velocity.

2. **The fundamental Wronskian.** Taking rectangular axes through O and letting x, y be the coördinates of P, the equations of motion are

* Received April 23, 1926, in revised form January 18, 1928. Presented to the American Mathematical Society, December 29, 1925.

$$x'' = -\frac{x}{r}\varphi, \qquad y'' = -\frac{y}{r}\varphi, \tag{1}$$

where

$$r^2 = x^2 + y^2. \tag{2}$$

Here, as in the sequel, primes denote derivatives with respect to the time t.

To avoid the consideration of a number of special cases we shall assume that φ is an analytic function of r and v and that φ does not vanish identically. Also, for simplicity of statement we shall include among the conics with focus at O the family of degenerate conics composed of straight lines through O.

Consider now the Wronskian of the three functions x', y', and r',

$$W = \begin{vmatrix} x' & y' & r' \\ x'' & y'' & r'' \\ x''' & y''' & r''' \end{vmatrix}, \tag{3}$$

the identical vanishing of which is a necessary and sufficient condition for the linear dependence of the functions.* We shall prove the following

THEOREM I. *The necessary and sufficient condition that the path be a conic with focus at the center of force is that* $W \equiv 0$.

Let the path be a conic with directrix

$$lx + my + p = 0 \tag{4}$$

and eccentricity e. Then, by the definition of a conic, the particle moves so that its distance from O is e times its distance from the directrix; that is

$$r = e\frac{lx + my + p}{\sqrt{l^2 + m^2}}, \tag{5}$$

---

* That $W \equiv 0$ is a necessary condition for a relation of the form

$$Ax' + By' + Cr' = 0,$$

A, B, and C being constants not all zero, is obvious, provided the third derivatives exist. That this condition is also sufficient has been proved (a) if the functions are analytic in t; or (b) if one of the Wronskians of two of the functions does not vanish at all. (Various sufficient conditions have been treated by Bôcher, Trans. Amer. Math. Soc. 2 (1901), 139–149.) Under our assumption that φ is analytic in r and v the coördinates are analytic in t, and (a) applies.

We could remove the requirement that φ be analytic and base the treatment on condition (b). We note subsequently that the Wronskian of x' and y' has the value

$$\begin{vmatrix} x' & y' \\ x'' & y'' \end{vmatrix} = \frac{\varphi}{r} D,$$

where D is a constant which does not vanish except for motion along a line through 0. If we require that φ should not vanish anywhere on the path, then (b) establishes the sufficiency of the condition.

where l and m are not both zero. The circle, which has no finite directrix, has the equation $r = c$; and the straight line through O has the equation $lx + my = 0$. These can all be put in the form

$$Ax + By + Cr + K = 0, \tag{6}$$

where A, B, and C are not all zero. Differentiating,

$$Ax' + By' + Cr' = 0. \tag{7}$$

It follows from this linear dependence that $W \equiv 0$.

Conversely, if $W \equiv 0$, a relation of the form (7) holds, where A, B, and C are not all zero. On integrating we get (6). If $C \neq 0$, we can write the equation in the form

$$r = lx + my + p.$$

If $l = m = 0$, we have the circle $r = p$. Otherwise we have

$$r = \sqrt{l^2 + m^2}\,\frac{lx + my + p}{\sqrt{l^2 + m^2}}, \tag{8}$$

a conic with focus at O, with directrix $lx + my + p = 0$, and with eccentricity $\sqrt{l^2 + m^2}$.

If $C = 0$, (6) is a straight line; and we find readily, by the use of (1), that $K = 0$, so the straight line passes through O. The condition is thus sufficient in all cases.

We next evaluate W. If we differentiate (2) three times, substitute the resulting values of r', r'', r''' into (3), and reduce by means of the relations (1), we find by direct algebraic manipulation

$$W = -\frac{D}{r^2}(r'' + \varphi)(2r'\varphi + r\varphi'),$$

where

$$D = \begin{vmatrix} x & y \\ x' & y' \end{vmatrix}. \tag{9}$$

Since

$$rr'' + r'^2 = xx'' + yy'' + x'^2 + y'^2 = -r\varphi + v^2,$$

and so

$$r'' + \varphi = \frac{v^2 - r'^2}{r} = \frac{D^2}{r^3},$$

we have

$$W = -\frac{D^3}{r^5}(2r'\varphi + r\varphi') = -\frac{D^3}{r^6}\frac{d}{dt}(r^2\varphi). \tag{10}$$

3. Determination of the acceleration. Since the force is central D is a constant and vanishes only if the motion is in a straight line. These

rectilinear paths through the origin exist whatever the law of acceleration may be. Excepting these orbits, the Wronskian vanishes if and only if*

$$\frac{d}{dt} r^2 \varphi = 0; \tag{11}$$

whence, integrating,

$$\varphi = C/r^2.$$

Hence, except for the straight line paths, the acceleration along each path is according to the Newtonian law. The constant C, however, may vary from path to path.

We now assume that φ depends upon the distance and the velocity alone,

$$\varphi = \varphi(r, s), \qquad s = v^2. \tag{12}$$

From (11) we have

$$2r\varphi r' + r^2\left(\frac{\partial \varphi}{\partial r} r' + \frac{\partial \varphi}{\partial s} s'\right) = 0.$$

Since

$$s' = 2(x'x'' + y'y'') = -2r'\varphi,$$

this may be written

$$rr'\left(2\varphi + r\frac{\partial \varphi}{\partial r} - 2r\varphi\frac{\partial \varphi}{\partial s}\right) = 0.$$

The factor $r' = 0$ yields circular orbits, $r = C$. The Wronskian vanishes for all orbits if and only if

$$r\frac{\partial \varphi}{\partial r} - 2r\varphi\frac{\partial \varphi}{\partial s} + 2\varphi = 0. \tag{13}$$

The partial differential equation (13) is readily integrated, with the result stated in the following theorem.

THEOREM II. *If φ is a function of r and v alone all orbits will be conics with one focus at the center if and only if φ, r, and v satisfy a functional relation of the form*

$$F(\varphi r^2,\ 2r\varphi - v^2) = 0. \tag{14}$$

4. **The orbit.** The constants l, m, and p, which determine the orbit, can be found by solving the equations

$$\begin{aligned} r &= lx + my + p, \\ r' &= lx' + my', \\ r'' &= lx'' + my''. \end{aligned}$$

* This equation can, of course, be arrived at by the usual methods,—without employing the Wronskian.

Reducing by means of (1) and (2), we find after an easy calculation

$$l = \frac{x}{r} - \frac{y' D}{\varphi r^2}, \qquad m = \frac{y}{r} + \frac{x' D}{\varphi r^2}, \qquad p = \frac{D^2}{\varphi r^2}.$$

The eccentricity e of the orbit is $\sqrt{l^2+m^2}$, whence

$$e^2 - 1 = l^2 + m^2 - 1 = \frac{D^2(v^2 - 2r\varphi)}{r^4}. \tag{15}$$

Here φr^2 and $v^2 - 2r\varphi$ are constant for any orbit. The path is an ellipse. a hyperbola, or a parabola according as $v^2 - 2r\varphi$ is negative, positive, or zero.

5. **The analogue of Kepler's third law.** If the orbit is an ellipse the semi-major axis is

$$a = \frac{p}{1-e^2} = \frac{\varphi r^2}{2r\varphi - v^2}, \tag{16}$$

and the area is

$$A = \pi a^2 \sqrt{1-e^2} = \frac{\pi \varphi r^2 D}{(2r\varphi - v^2)^{3/2}}.$$

Then the period is

$$T = \frac{2A}{D} = \frac{2\pi \varphi r^2}{(2r\varphi - v^2)^{3/2}}. \tag{17}$$

From these we have

$$\varphi r^2 = \frac{4\pi^2 a^3}{T^2}, \qquad 2r\varphi - v^2 = \frac{4\pi^2 a^2}{T^2}.$$

Then (14) may be written

$$F\left(\frac{4\pi^2 a^3}{T^2}, \frac{4\pi^2 a^2}{T^2}\right) = 0. \tag{18}$$

This relation between the period and the semi-major axis holds for all elliptic orbits; it is the analogue of Kepler's third law.

THEOREM III. *Any functional relation*

$$f(a, T) = 0 \tag{19}$$

may be assumed as the analogue of Kepler's third law and the acceleration is thereby determined.

For, using (16) and (17), we can write (19) in the form

$$f\left(\frac{\varphi r^2}{2r\varphi - v^2}, \frac{2\pi \varphi r^2}{(2r\varphi - v^2)^{3/2}}\right) = 0. \tag{20}$$

This is a relation of the form (14), and φ is determined.

In particular, if we assume the usual third law, $T^2 = Ka^3$, there results the Newtonian case, $\varphi = K'/r^2$.

6. **Examples.** In the following examples we take certain simple forms of the relation (14) and point out a few of the properties of the motions.

$$\text{I.} \quad 2r\varphi - v^2 = 0, \quad \text{or} \quad \varphi = \frac{v^2}{2r}.$$

We see from (15) that the orbits are all parabolas.

$$\text{II.} \quad 2r\varphi - v^2 = k^2, \quad \text{or} \quad \varphi = \frac{v^2 + k^2}{2r}.$$

All orbits are ellipses. The analogue of Kepler's third law, (18), is

$$\frac{4\pi^2 a^3}{T^2} = k^2, \quad \text{or} \quad T = \frac{2\pi}{k} a.$$

The period is proportional to the semi-major axis.

III. Find φ so that the period is the same for all elliptical orbits. From (17),

$$\frac{4\pi^2 \varphi^2 r^4}{(2r\varphi - v^2)^3} = c^2.$$

This gives φ in implicit form.

IV. An interesting case is that in which all elliptic orbits have the same major axis. From (16),

$$\frac{\varphi r^2}{2r\varphi - v^2} = a, \quad \text{or} \quad \varphi = \frac{a v^2}{r(2a - r)}.$$

A body projected from any point within a distance $2a$ of the origin and with any velocity will move in an ellipse of semi-major axis a.

V. An interesting proposition of Celestial Mechanics is the following. Let particles be projected simultaneously from a point with the same velocity, but with varying directions, the velocity being such that the orbits are ellipses. Then all orbits have the same mean distance and hence the same period. The particles will, therefore, subsequently return to the point of projection simultaneously.

This proposition is true for the more general accelerations treated in this paper. At the moment of projection r and v are the same for all the particles; hence, φ is the same. Hence, from (16) and (17), a and T are the same for all the orbits.

THE RICE INSTITUTE.